HOKKAIDO UNIVERSITY

Title	A Study on Emoticon Recommendation System Considering the Content of Text Messages
Author(s)	卜部, 有記
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第14593号
Issue Date	2021-03-25
DOI	https://doi.org/10.14943/doctoral.k14593
Doc URL	https://hdl.handle.net/2115/81464
Type	doctoral thesis
File Information	Yuki_Urabe.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Doctoral dissertation

博士論文

A Study on Emoticon Recommendation System
Considering the Content of Text Messages

(テキストメッセージの内容に対して適切な顔文字を推薦する
顔文字推薦システムに関する研究)

Hokkaido University
Graduate School of Information Science and Technology

Yuki Urabe

北海道大学
大学院情報科学研究院

卜部　有記

学位論文の要旨

近年，インターネット上では Twitter や Facebook などのソーシャルメディア上でテキストベースのコミュニケーションが活発である．テキストによるコミュニケーションは，相手の顔が見えないことから気持ちや意図の伝達が困難である．このことからユーザは，顔文字（「(・▽・)」）をテキストメッセージ（以下，「メッセージ」は「テキストメッセージ」を指す）中に挿入することで相手に気持ちや意図を明確に伝える傾向にある．しかし，市中の顔文字入力アプリでは 100 万種類，また，iOS や Android の日本語キーボード上では約 300 種類以上の顔文字が存在するため，ユーザのメッセージの内容に沿って適切な顔文字を登録されている顔文字の中から探し出して挿入することは困難である．現状は，過去に使用した顔文字を容易に見つけられるように提示する以外の支援はなく，ユーザのメッセージの内容を考慮して適切な顔文字を提示する支援は行われていない．そのため，ユーザは使用したことのない顔文字を挿入したい場合は，顔文字を探さなければならない．

メッセージ中の顔文字の使われ方に関する研究では，メッセージ中の顔文字の役割は，感情の強調や感情の明確化であることが明らかにされている．また，顔文字自体が表す感情表現に着目して顔文字がどの感情をどの程度表すかという観点から顔文字データベースを構築する方法が過去に提案された．他にも，メッセージの内容に対して適切な顔文字を自動的に挿入あるいは推薦することを目的とした研究がある．具体的には，感情を表す単語（以下、「感情語」とする）と顔文字を対応づけた辞書を用意し，メッセージ中に感情語が存在する場合に対応づけられた顔文字を自動的に挿入するという手法がある．また，メッセージ中の単語の特徴から感情，コミュニケーション，動作などのカテゴリを予測し，予測したカテゴリにおいて利用頻度の高い顔文字を提示するという手法が過去に提案された．しかし，1 つ目の手法に関しては，顔文字は同一の感情を表すものが多様に存在することからメッセージが表す感情に合わせて毎回同じ顔文字が挿入されるとは限らず，多様な顔文字が挿入される可能性がある．そのため，メッセージが表す感情に対して特定の顔文字しか挿入できない手法はユーザにとって不十分だと考えられる．また，2 つ目の手法に関しては，各カテゴリにおける顔文字の使用傾向およびメッセージがどのカテゴリに該当するかをアノテーションしたコーパスを作成する手間がかかる．加えて，これら 2 つの手法は，ユーザの好みについては考慮していないため，メッセージの内容に対応した顔文字は推薦できるが，それがユーザの好みに適合した顔文字であるとは限らない．私は，まず前述の 2 つの手法に対して述べた問題を解決するため，メッセージ中の感情表現に基づいて感情タイプを予測し，予測結果を用いて顔文字の感情強度とユーザの顔文字の選択履歴に基づいて適切な順番に顔文字を推薦する手法を提案した．この手法により，メッセージの内容とユーザの好みを考慮して顔文字を推薦することが可能となる．次に，前述の既存手法はメッセージ中に感情表現など特徴的な単語が存在しない場合は顔文字を推薦できないことに着目した．そして，メッセージ中に感情表現を含まない場合においても適切に顔文字を推薦するために，インターネット上の自然言語文を学習した事前学習済みモデルを利用して，メッセージと一緒に挿入された顔文字を学習することで表層的には表れないメッセージの内容を全体的に考慮して顔文字を推薦する手法を提案した．

本学位論文は 5 章からなる．第 1 章では本学位論文の背景と目的，そして本研究による貢献を述べる．第 2 章では，顔文字のテキストメッセージ中の役割や顔文字挿入支援手法，さらに絵文字にも着目して広く関連研究を述べる．第 3 章では，1 つ目の提案手法である，メッセージ中の感情表現と一致する感情の強度順，およびユーザの選択履歴に基づいて顔文字を推薦する顔文字推薦システム（EBReS-η : Emoticon-prioritized History-Based Emoticon Recommendation System）について述べる．提案システムを実現するにあたり，59 個の各顔文字に対して 10 種類の感情の強度を人手により評価を行い，顔文字と感情強度によるデータベースを作成した．提案システムの評価は，ユーザ実験を通じて行った．また，提案手法の有効性を明らかにするために，iOS で利用されている選択履歴による方法と提案手法とを比較することにより評価した．評

価の結果，提案手法の方が既存手法よりもユーザが 5 位以内に提示された顔文字の中から選択した割合が 42.3 ポイント高いことが明らかとなり，本手法の有効性が確認できた．また，SD（Semantic Differential）法によりユーザによる各手法への印象（使いやすさ，便利さ，など）を評価したところ，提案手法の方が平均 1.30 ポイント高いことが確認でき，ユーザのシステムに対する印象が既存手法よりも良いことが明らかとなった．本論文には，上位に提示した顔文字が選択されなかった事例について原因分析を行った結果も示す．

第 4 章では，2 つ目の提案手法として，メッセージ中に感情表現を含まない場合においても，日本語文の事前学習済みモデルを利用してメッセージと顔文字の使われ方を学習することにより，メッセージの全体内容に対して適切な顔文字を推薦するシステム（Deep-ERS：Deep-Emoticon Recommendation System）について述べる．提案システムの評価は，第 3 章で述べた既存手法，および従来の機械学習手法に用いられてきたメッセージの表層情報を特徴量として学習する方法を用いて顔文字の提示順位の比較を行った．テキストメッセージと 167 種類の顔文字のどれか 1 つを含む，合計 37,622 件のテストデータ（テキストメッセージと顔文字のセット）を用いて評価した結果，提案手法の方が既存手法よりも正解の顔文字が 25 位以内に出現した割合が最大で 23.5 ポイント高いことが確認でき，本手法の有効性を明らかにした．本論文には，成功事例の他，正解の顔文字が上位に提示されなかったケースを抽出して原因分析を行った結果も示す．

最後に第 5 章では，本研究を総括し結論を述べた．本研究により，メッセージ中の感情表現に対して顔文字の感情強度とユーザの選択履歴に基づいて顔文字を推薦することで，メッセージの内容かつユーザの好みに合った顔文字の提示を実現した．さらに，メッセージ中に感情表現が存在しない場合も事前学習済みモデルを利用してメッセージと顔文字の使われ方の特徴を学習して顔文字を推薦することでメッセージ中の感情表現の有無に関わらずに適切な顔文字を提示することを可能にした．

Contents

1 Introduction **12**

1.1 Background . . . 12

1.2 Contribution of this research . . . 14

2 Related research **17**

2.1 Role of emoticons in text messages . . . 17

2.2 Research on emoji . . . 17

2.3 Research on emoticons . . . 19

3 Emotion-prioritized History-Based Emoticon Recommendation System (EBReS-η) **20**

3.1 Introduction . . . 20

3.2 Procedure of EBReS-η . . . 20

3.2.1 Affect analysis system: ML-Ask . . . 21

3.2.2 Emoticon database . . . 22

Survey . . . 23

Survey results . . . 23

3.3 Evaluation experiment . . . 24

3.3.1 Participants . . . 25

3.3.2 Emoticon recommendation application . . . 26

3.3.3 Experiment procedure . . . 26

3.3.4 Evaluation measurements . . . 28

(A) Average rank of the chosen emoticons . . . 28

(B) Proportion of chosen emoticons appeared within the top five . . . 28

(C) Users' impression towards the systems . . . 28

(D) User vote for systems' performance and preference . . . 30

3.4 Results and discussion . . . 30

3.4.1 (A) Average rank & (B) proportion of emoticons chosen within the top five . . . 30

3.4.2 (C) Participants Impression using SD Scale . . . 30

3.4.3 (D) User vote for system's performance and preference . . . 34

3.4.4 Error analysis . . . 34

3.4.5 Discussion . . . 37

3.5 Conclusion of Chapter **3** . . . 38

4 Deep Emoticon Recommendation System (Deep-ERS) 39
4.1 Introduction . . . 39
4.2 Preliminary investigation . . . 39
4.2.1 Data description . . . 39
Annotation of polarity and emotion types to sentences . . . 40
Emotion database . . . 40
4.2.2 Results . . . 41
4.3 Deep-ERS . . . 45
4.3.1 Pre-trained models . . . 46
Word2vec . . . 46
ELMo . . . 47
BERT . . . 47
4.3.2 Prediction models . . . 48
Word2vec+SVM . . . 48
Word2vec/ELMo+BiLSTM+Attention . . . 48
BERT+fine-tuning . . . 50
4.4 Evaluation experiment . . . 51
4.4.1 Evaluation methods . . . 51
(A) Recommendation accuracy . . . 51
(B) Mean Average Precision (MAP) . . . 52
4.4.2 Data description . . . 52
4.4.3 Baseline methods . . . 53
Frequency method . . . 55
BoW+SVM method . . . 55
TF-IDF+SVM method . . . 56
WordCharGram+SVM method . . . 56
4.4.4 Implementation for each proposed prediction model . . . 57
Word2vec+SVM . . . 57
Word2vec+BiLSTM+Attention . . . 58
ELMo+BiLSTM+Attention . . . 59
BERT+fine-tuning . . . 59
4.5 Results and discussion . . . 59
4.5.1 (A) Recommendation accuracy . . . 59
4.5.2 (B) MAP . . . 61
4.5.3 Comparison of the prediction models . . . 63
4.5.4 Problem of recommending lower rank emoticons . . . 66
4.6 Discussion . . . 69
4.7 Conclusion of Chapter **4** . . . 70

5 Conclusion 71

A Appendix for Chapter 3 **73**

A.1 Emoticon database (59 emoticon registered) 73

A.2 Sentences used in the experiment 73

A.3 Factor analysis 73

A.4 Factor analysis result 74

B Appendix for Chapter 4 **85**

B.1 Emoticon database (299 emoticon registered) 85

List of Figures

1.1 Two examples of standard input method for emoticons. The left figure shows a word converting method using IME. When "かおもじ" (*kaomoji*, meaning "facemarks" in Japanese) is typed, a list of emoticons appear on the display. The right figure shows an emoticon selection method when iOS Japanese keyboard is used. 13
1.2 Examples of emoticons comprised of special unicode symbols [1] 14

2.1 Examples of emojis in text messages 18
2.2 Visualisation of 100 emojis vectors, reduced to two dimensions with t-SNE [2] 18
2.3 Emojis used in SemEval 2018 Task 2, listed in frequency order [3] 18

3.1 System procedure of **EBReS-η** 21
3.2 Example of the output given by ML-Ask [4] 22
3.3 Example of emoticon ratings in each of the 10 emotions [5] 23
3.4 Example of an emoticon connected with average values of each of the ten emotions. The red bars are values larger than 3.00 and the blue bars are values smaller than 3.00. 23
3.5 Number of emoticons rated more than 3.0 for each emotion type 25
3.6 Emoticon recommendation procedure in the application 27
3.7 Example of a vote for system's performance and users' preferences 30
3.8 Number of chosen emoticons by rank. "Pro" stands for "Proportion of emoticons chosen within the top five". "Ave" stands for "Average ranking of the chosen emoticons". Red line indicates the average ranking. 31
3.9 Participants impression of each system using 22 impression word pairs 32
3.10 Biplot of the two factor models for factor 1 (Usefulness) and 2 (Novelty). Factor 1 explains 42.0% and factor 2 explains 12.0% of the total variance. Red star = **EBReS-η**, green triangle = **EBReS**, and gray cross = **HiBReS**. Blue dots indicate abbreviations for SD adjectives: Amu = Amusement, Imp = Impressiveness, Fri, Friendliness, Han = Handiness, Sp = Speed, Con = Convenience, Nec = Necessity, Lig = Lightness, Rec = Recognizability, Fon = Fondness, Nov = Novelty, Sim = Simplicity, Int = Interest, Fre = Freshness, Acc = Accurate, Use = Useful, Viz = Visibility, Eas = Easiness, Cho = Choosability, Spe = Specialness, Sma = Smartness, Sat = Satisfactory. 35

3.11 Biplot of the two factor models for factor 3 (Simplicity) and 4 (Process quality). Factor 3 explains 5.9% and factor 4 explains 5.6% of the total variance. Red star = **EBReS-η**, green triangle = **EBReS**, and gray cross = **HiBReS**. Blue dots indicate abbreviations for SD adjectives: Amu = Amusement, Imp = Impressiveness, Fri, Friendliness, Han = Handiness, Sp = Speed, Con = Convenience, Nec = Necessity, Lig = Lightness, Rec = Recognizability, Fon = Fondness, Nov = Novelty, Sim = Simplicity, Int = Interest, Fre = Freshness, Acc = Accurate, Use = Useful, Viz = Visibility, Eas = Easiness, Cho = Choosability, Spe = Specialness, Sma = Smartness, Sat = Satisfactory. 36

4.1 Percentage of sentences with emoticons containing emotive phrases [6]. "Agree" indicates that polarity type of an emotive phrase matches emoticon's polarity type within a given sentence. 41
4.2 Ratios of emoticons' emotion types included in sentences categorized by emotion [6]. The number of sentences for each emotion type is shown in the table and their percentage in brackets. 43
4.3 System procedure 44
4.4 Word2vec model [6] 46
4.5 ELMo model [6] 47
4.6 Outline of **Word2vec+SVM** model [6] 48
4.7 **Word2vec/ELMo+BiLSTM+Attention** model [6] 49
4.8 **BERT+fine-tuning** model [6] 50
4.9 BERT input representation [7] 51
4.10 Procedure for creating training/test data samples [6] 54
4.11 Examples of creating Bag of Words feature vectors from sentences 56
4.12 An example of creating a feature vector for a sentence with **WordCharGram** 57
4.13 Number of unique emoticons that appeared with the emotive expression in sentences in the training data [6]. The top two lowest bars (ANGER and SHYNESS) are emphasized with stripes to mirror the discussion written in 4.5.1. 62
4.14 Distribution of average precision for polarity [6]. Note that sentences containing emoticons neither expressing POSITIVE nor NEGATIVE are excluded (see 4.4.1). Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 65
4.15 Distribution of average precision for emotion [6]. Note that sentences containing emoticons neither expressing POSITIVE nor NEGATIVE are excluded (see 4.4.1). Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 65

4.16 Average rank per emoticon of each prediction model (represented as markers) and data size of each emoticon in the training (blue line) and test (orange line) data [6]. Different colors mark results of following algorithms: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 67

A.1 Factor analysis 73

A.2 Correlation of 22 bipolar words and eigenvalues. Four values in the red box are larger than 1 which is a determination of the number of factors. 83

A.3 Factor analysis results when the number of factors is 4 84

List of Tables

1.1 Vertical and horizontal type emoticons . . . 13
1.2 Aspect consideration for each method . . . 15

3.1 Correlation coefficient of ten emotion types calculated from the average ratings of each emoticon. ** indicates the P value lower than 0.05 and * indicates the P value lower than 0.01. P value lower than 0.05 means that the correlation is statistically significant. . . . 24
3.2 Details of participants regarding their device types used to communicate and their daily message and emoticon usage . . . 26
3.3 Example of sentences shown to participants. Emotive expressions are highlighted in bold. 28
3.4 22 impression word pairs . . . 29
3.5 Factor loadings of each of the 22 impression word pairs in the semantic differential (SD) scale . . . 33
3.6 Rankings on the systems' performance and users' preferences. The highest value for each method is emphasized with bold. . . . 34
3.7 Number of sentences in which the emotion type of selected emoticons and emotive words in sentences disagreed. POSITIVE, NEGATIVE, and OTHER are the polarity types of the emotive phrase recognized within sentences. . . . 37
3.8 Examples of sentences in which the polarity types of emotive phrase and the selected emoticon disagreed with each other. Emotive phrase in the sentence is emphasized with bold. . . . 37

4.1 Examples of sentence-emoticon pairs. Emotive words are emphasized in bold. . . . 40
4.2 Example emoticons from emoticon dictionary introduced in [8]. The highest value for each emoticon is emphasized in bold. . . . 41
4.3 167 emoticons and their emotion and polarity types . . . 42
4.4 Examples of sentence-emoticon pairs that their emotion types do not match with each other. Emotive words are emphasized in bold. . . . 44
4.5 Number of sentences for each emotion type (training data) . . . 53
4.6 Number of sentences for each emotion type (test data) . . . 53
4.7 The best parameters for each baseline method . . . 57
4.8 The best parameters for each prediction model (models using SVM) and number of dimensions of pre-trained models. . . . 57
4.9 The best parameters for each prediction model (models using deep learning techniques) . 58

4.10 Recommendation accuracy (A@25) (%). Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 60
4.11 P value of Wilcoxson signed-rank test between each method. "*" for values higher than 0.05. Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 61
4.12 Examples of sentences containing the emotive word, "怒る" (*okoru*, "angry"). 62
4.13 Examples of sentences containing the emotive word, "恥ずかしい" (*hazukashii*, "embarrassed"). 63
4.14 MAP of each prediction model (point). "Average" means the average MAP of the entire samples. The highest values are highlighted in bold. Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 64
4.15 Ranking of the prediction models based on the results of recommendation accuracy and MAP 66
4.16 MAP of each prediction model for emoticons less than 200 samples in the test data. The highest values are highlighted in bold. Numbers stand for the following methods: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**. 68
4.17 Average proportion of emoticons of the same polarity/emotion type used in similar sentences in the training data 69
4.18 Example of a sentence from training data where emoticon usage is different from the samples in the test data (the first sentence in bold is the source for finding similar sentences) 69

A.1 Emoticon database (from ID 1 to ID 30) described in Chapter **3**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 75
A.2 Emoticon database (from ID 31 to ID 59) described in Chapter **3**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 76
A.3 Sentences used in Group 1. Emotive phrases are emphasized with bold. 77
A.4 Sentences used in Group 2. Emotive phrases are emphasized with bold. 77
A.5 Sentences used in Group 3. Emotive phrases are emphasized with bold. 78
A.6 Sentences used in Group 4. Emotive phrases are emphasized with bold. 78

A.7 Sentences used in Group 5. Emotive phrases are emphasized with bold. 79
A.8 Sentences used in Group 6. Emotive phrases are emphasized with bold. 79
A.9 Ratings of HiBReS system evaluated by 30 participants. 80
A.10 Ratings of EBReS system evaluated by 30 participants. 81
A.11 Ratings of EBReS-η system evaluated by 30 participants. 82

B.1 Emoticon database (from ID 60 to ID 90) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 86
B.2 Emoticon database (from ID 91 to ID 120) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 87
B.3 Emoticon database (from ID 121 to ID 150) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 88
B.4 Emoticon database (from ID 151 to ID 180) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 89
B.5 Emoticon database (from ID 181 to ID 210) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 90
B.6 Emoticon database (from ID 211 to ID 240) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 91
B.7 Emoticon database (from ID 241 to ID 270) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 92
B.8 Emoticon database (from ID 271 to ID 299) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS. 93

Chapter 1

Introduction

1.1 Background

Computer-Mediated Communication (CMC) has become popular in recent years as it allows people to communicate regardless of time, place, and familiarity (i.e. whether or not they know each other). However, in contrast to Face-to-Face communication, text-based communication in CMC lacks the ability to convey nonverbal cues such as facial expression, tone of voice, and body movement [9]. These nonverbal cues help humans to understand feelings and intentions of interlocutors [10] and may take up an estimated 93% of communication (38% from voice and 55% from facial expression) [11]. To prevent confusion and express users' intentions accurately in CMC, it is necessary to find a way to compensate for this lack of nonverbal cues.

Emoticons, which are marks expressing faces or movements composed of letters and symbols, may serve as nonverbal surrogates in CMC [12, 13, 14]. They are widely used in CMC, namely in instant messaging and social media to express one's feelings, enhance the emotion of a sentence, and express humor [15, 16, 17]. Recipients can understand the sender's intended feelings better with emoticons included in a sentence than by receiving only words [18, 19, 20]. Information conveyed by emoticons has great importance in CMC which should not be ignored, and thus, research of emotion analysis including emoticons, and the development of interfaces that support users expressing their feelings using emoticons are highly important.

Emoticons can be divided into two types: vertical style (Table 1.1, see "Vertical"), mainly used in the Western world, and horizontal style (Table 1.1, see "Horizontal"), mainly used in the Far East [21]. Vertical style emoticons are composed of symbols and Latin letters and are rotated by 90 degrees. The horizontal style is composed of symbols from other writing systems like the Cyrillic alphabet, Chinese ideograms, or Japanese and Korean syllabaries [21, 22, 23, 24]. Moreover, the horizontal style is not rotated which makes it more easily comprehensible to a reader [21]. Since the horizontal style consists of a larger number of symbols and letters than the vertical one, over one million examples of the latter were discovered on the Internet[1], while only around 300 vertical-style emoticons are listed in Wikipedia[2]. Moreover, like in the example in Table 1.1, there are various emoticons that express certain emotion

[1] Available in Flick (`https://kaomoji.tokyo/`), an input method application with a dictionary comprising over 2 million emoticons.

[2] List of emoticons: `https://en.wikipedia.org/wiki/List_of_emoticons`

Table 1.1: Vertical and horizontal type emoticons

Emotion	Vertical	Horizontal
Joy	:) :D :-)	(^_^) (^▽^) (*´▽`*)
Sadness	:(:-c :-<	(T_T) (ToT) (;_;) (/_;)
Amazement	:O :-o	(￣□￣;) (;゜0゜) ! (◎_◎;)

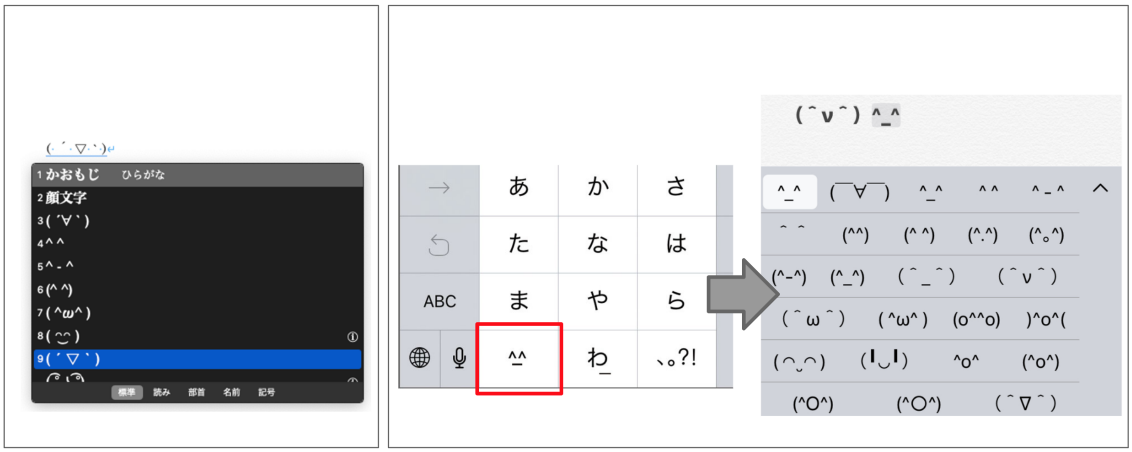


Figure 1.1: Two examples of standard input method for emoticons. The left figure shows a word converting method using IME. When "かおもじ" (*kaomoji*, meaning "facemarks" in Japanese) is typed, a list of emoticons appear on the display. The right figure shows an emoticon selection method when iOS Japanese keyboard is used.

category (e.g. (^_^), (^▽^), and (*´▽`*) express "Joy"). These large numbers of emoticons are sophisticated enough to express nuances in meaning and may richen the quality of communication in CMC. However, it is difficult for users to find appropriate emoticons to express their intentions from these large number of emoticons and from various emoticons that express the same emotion type.

There are four different input methods for inputting emoticons that are discovered through user observation [25].

- Manual input
- Convert a word to an emoticon
- Copy and paste an emoticon from webpages or mobile applications where emoticons are registered
- Select an emoticon from an emoticon dictionary

At the beginning of the 80s', emoticons were inputted manually from computer keyboards [26, 27]. Simple emoticons (e.g. ":-)" or "(^_^)") were easy to insert by typing symbols directly from the keyboard. However, many emoticons of those comprised of symbols that users scarcely know have appeared on the Internet which increased the difficulty of typing them. The second input method is converting a word to an emoticon using an input method editor (IME) which is commonly used when communicating in certain languages such as Japanese, Chinese, and Korean (Figure 1.1, left). This method has made inserting emoticons easier for users because they only need to type the relevant word (e.g. "かおもじ"

Figure 1.2: Examples of emoticons comprised of special unicode symbols [1]

(*kaomoji*[3], meaning "facemarks") or "嬉しい" (*ureshii*, meaning "Joy"), etc.) and convert into expected emoticons. Some emoticons are registered by default, however, users have to register word-emoticon pair inside their keyboard settings and also remember which word is linked to a given emoticon. The third input method is copying and pasting an emoticon from webpages or mobile applications where a list of emoticons is registered. Recently, many unique emoticons comprised of special Unicode symbols have appeared on the Internet and are popular among young generations (Figure 1.2). These emoticons cannot be input directly from computer and tablet keyboards and often appear on webpages and dedicated mobile applications [1, 28]. This enables users to use these special emoticons, however, users may feel troublesome to open a webpage or application to search and copy the desired emoticon and then paste it on to the text. The fourth method is to select an emoticon from an emoticon dictionary. This method is very common for smartphone and tablet users in Japan because there is an emoticon button to show a list of emoticons on the Japanese keyboard (Figure 1.1, right). The number of emoticons registered on the keyboard on smartphones and tablets are 343 in iOS 13 and 386 in Android 10 and it has been continually increasing from the previous OS versions which included around 200 emoticons. Even if they represent only a small fraction of all existing emoticons, these numbers are still large and the emoticons' expressions are diverse, thus it is difficult for users to find an emoticon that they consider appropriate for their messages. The current function that assists users with selecting an emoticon is designed to display a set of previously chosen ones on the top of the list and does not take the content of a message into account when recommending. However, people choose emoticons depending on what kind of information they want to pass, for instance, positive emoticons are likely to be chosen in positive contexts, and vice versa for negative ones [20, 29]. Therefore, I conducted a study and proposed emoticon recommendation systems that assist users' emoticon selection by recommending emoticons that match the content of their input and their preference for emoticons.

1.2 Contribution of this research

To realize a system that can recommend more appropriate and desired emoticons to users' inputs, I believe that there are two aspects that need to be considered. These aspects are described below.

Aspect (1) Considering the content of the input. Users select emoticons regarding the content of their text messages (see "Related research" in 2.1). Therefore, I believe that it is important for a system to recommend emoticons based on the content of the input so that users can obtain desired emoticons for their input easily.

[3]Italic font indicates foreign language transcripts.

Table 1.2: Aspect consideration for each method

Aspect \ Method		Current iOS/Android	Suzuki *et al.* [30]	Emura *et al.* [31]	EBReS-η [5] (Proposed method 1)	Deep-ERS [6] (Proposed method 2)
(1) Considering the content of the input	A. when emotive phrase is contained	✗	✓	✓	✓	✓
	B. when emotive phrase is **not** contained	✗	✗	✗	✗	✓
(2) Considering users' emoticon preference		✓	✗	✗	✓	✗

Aspect (2) Considering users' emoticon preference. There are a variety of emoticons expressing the same emotion category as explained in 1.1. From this fact, I hypothesize that users select emoticons not only from the content of the input message but also based on their personal preferences for emoticons. If my hypothesis is correct, I believe that recommending emoticons considering users' preference is also important to improve the easiness of selecting emoticons for users.

Table 1.2 describes how the current methods (the method used in iOS and Android, Suzuki *et al.*'s method [30], and Emura *et al.*'s method [31]) consider these aforementioned aspects. Also, I added my two proposed methods in the table for comparison with the current methods. The summary of my proposed method is explained in the next paragraph and the details are given in Chapters 3 and 4. The ✓ in the table means that the method takes the aspect into account for recommendation and ✗ indicates it does not. For Aspect (1-A), the method used in iOS and Android does not analyze the input for recommendation while Suzuki's method [30] recommends an emoticon that is linked to the emotive phrase contained in the input, and Emura's method [31] recommends emoticons based on the emotion category of the input. However, none of these current methods and my first proposed method can recommend emoticons when no emotive phrase is present in the input message (Aspect (1-B)). For Aspect (2), the current method used in iOS and Android shows emoticons that were previously inserted in users' input sentences, therefore, it can be considered that this method can display emoticons regarding users' emoticon preference. On the other hand, Suzuki and Emura's methods do not consider users' preferences toward the emoticons.

The contribution of this research considering the aforementioned aspects is as follows:

- Proposed **Emoticon-prioritized History-Based Emoticon Recommendation System**[4] (**EBReS-η**) that reorders emoticons in the database based on users' emotive statements and their past selections (consideration of Aspects (1-A), (2)). The emoticon database containing 59 emoticons is annotated with the strength for each of ten emotion types by conducting a questionnaire. The user experiment showed that 72.3% of chosen emoticons were among the top five recommended by the proposed system, which is an improvement of 42.3 points over the method used in current smartphones, which is based only on users' past emoticon selections. Details are given in Chapter **3**.

- Proposed **Deep-Emoticon Recommendation System** (**Deep-ERS**) that takes the whole input message into account to recommend emoticons (consideration of Aspects (1-A), (1-B)). The system works for both input sentences that include or do not include emotive expressions with a help of a pre-trained model incorporated into the recommender system. The pre-trained model was created by learning large Japanese text data collected from the Internet. The evaluation experiment showed that this proposed method improved the baseline methods by a maximum of 23.5 points for recommendation accuracy. The baseline methods I used for comparison are the one using only users'

[4]Bold font in the sentence indicates proposed and baseline system and method names.

past emoticon selections (similar to Emura *et al.*'s method [31]) and other machine learning-based methods that learned the surface patterns of messages as features (e.g. Bag-of-words, word and character n-grams) and emoticons used with the messages. Details are given in Chapter **4**. This research also studied what kind of data the pre-trained model should learn from to recommend more suitable emoticons for input messages that are used in casual communication (e.g. communication in SNS).

Chapter 2

Related research

2.1 Role of emoticons in text messages

Several studies analyzed how emoticons are used in text messages. Emoticons are used to express emotion, appearing more often in casual communication with friends than in communication with strangers [20, 32, 33]. In messages, the same valence of emoticons tends to be used to enhance the emotion [15, 20]. On the other hand, it was discovered that positive emoticons are used in a negative context to soften the negative meaning, such as rejections, requests, or complaints [16, 17, 34]. Also, when emoticons are used in a neutral context, they serve as an emotional indicator [16]. Finally, emoticons are also used in sentences to express humor and sarcasm [32, 35].

2.2 Research on emoji

Emojis are also popularly used in CMC. They are used to express emotion which is similar to the role of emoticons (e.g. Figure 2.1 (1), inserting a smiley at the end of a sentence), used in a sentence as a substitution of words (e.g. Figure 2.1 (2), replacing the word "Japan" with an emoji of a Japanese flag), and used in a sentence to enhance the meaning (Figure 2.1 (3), adding an emoji of a car after the word "car"). In the field of natural language processing, Barbieri *et al.* [2] studied emoji semantics by constructing a skip-gram neural embedding model using both words and emojis extracted from Twitter based on the aforementioned role of emojis. They confirmed that emojis expressing similar meanings have similar vectors, which indicates that they are used similarly in texts. Figure 2.2, borrowed from [2], shows that emojis of similar meaning are plotted close to each other. For instance, emojis related to music such as a guitar, a piano keyboard, and a headset are shown closer to each other on the right side.

In the SemEval 2018 Task on Multilingual Emoji Prediction [3], participants were asked to predict emojis that are most likely to be used along with the text in English and Spanish tweets. The dataset used in this challenge comprised of 550K tweets (500K tweets for training data and 50K tweets for test data), each containing one of the 20 most frequently used emojis. The top F-scores of 35.99% and 22.36% were achieved by Çöltekin *et al.*'s method [36] for English and Spanish, respectively. Çöltekin *et al.*'s method used bag-of-N-grams features of both words and characters with Support Vector Machine (SVM) classifier. Zhang *et al.* [37] outperformed Çöltekin *et al.*'s F-score by 2.53 points for English tweets by fine-tuning the training data with BERT [7]. In their research, they stated that applying knowledge learned from

(1) I'm so happy to see you 😊

(2) I went to 🇯🇵 for sightseeing!

(3) I am going by car 🚗!

Figure 2.1: Examples of emojis in text messages

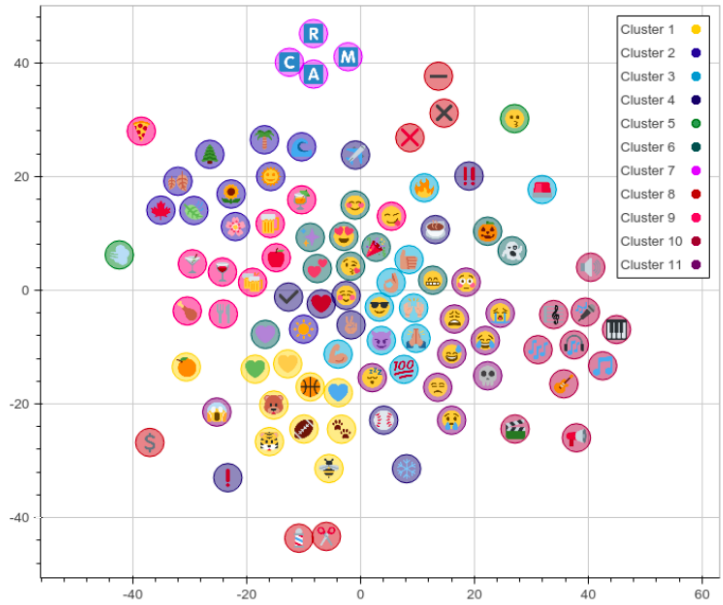


Figure 2.2: Visualisation of 100 emojis vectors, reduced to two dimensions with t-SNE [2]

Figure 2.3: Emojis used in SemEval 2018 Task 2, listed in frequency order [3]

other natural language processing tasks improves the results because such a model can utilize semantic information (i.e. the relation between words and sentences) in the new task. As mentioned before, the difference between emojis and emoticons (facemarks) is that the former expresses not only emotion with facial expressions but also objects, states, and so on. The latter, on the other hand, mostly express emotions with various kinds of facial appearance. As the emoji dataset introduced in [3] included both faces and objects (Figure 2.3) and the F-score of facial emoji prediction results tend to be lower than that of objects by the average of 23.65 points for English and 1.57 points for Spanish, it becomes quite obvious that recommending emoticons for sentences by applying machine learning techniques is a challenging task.

2.3 Research on emoticons

In the field of sentiment analysis, emoticons have gotten attention and have been considered useful for analyzing the sentiment of text messages. Yamamoto *et al.* [16] and Hogenboom *et al.* [38] proposed methods that included emoticon sentiment analysis to predict the sentiment of the whole sentence. Ptaszynski *et al.* proposed a method to analyze an emotion type of an emoticon-based on the theory of kinesics [39].

Rule-based emoticon recommendation methods were proposed by Suzuki *et al.* [30] and Emura *et al.* [31]. Suzuki *et al.* proposed a method that analyzes an emotive load in an input sentence and recommends emoticons based on the emotion type of the emotive expression. In their research, they used two lexicons: one containing emotion-expressing phrases categorized into emotion types, and one comprised of emoticons accompanied by emotive phrases. The emoticon lexicon was created through a questionnaire that aimed to choose a representative emoticon for each emotive phrase. Emura *et al.* proposed a method that recommends emoticons not only from the emotive phrases found in an input sentence but also from communication-related (e.g. greetings, congratulating, etc.) and activity-related (e.g. sleeping) expressions. In their research, they constructed a corpus of 3,975 sentences that contains emotional phrases, communication-related and activity-related expressions. The expression types (emotion, communication, and movement) were annotated by humans. Emoticon dictionary was also created by analyzing the corpus and registering a set of emoticons, an expression type, and an emoticon occurrence in the corpus. Their method recommended emoticons by predicting the expression type from the input by extracting similar texts from the corpus using the k-nearest neighbors algorithm and then showed frequently used emoticons that appeared in corpus sentences of the same expression type.

Chapter 3

Emotion-prioritized History-Based Emoticon Recommendation System (EBReS-η)

3.1 Introduction

This chapter describes the first proposed system **EBReS-η** (**Emotion-prioritized History-Based Emoticon Recommendation System**) that takes two aspects[1] into account to recommend more desired emoticons for users' input messages. This system recommends emoticons in a suitable order to an input made by users and enables them to easily choose an emoticon from a list of candidates. The distinctive feature of this system is that it recommends emoticons by analyzing emotive phrases in the input and then reorder an emoticon database to the order of emoticons that express the same emotion type of emotive phrases. Moreover, the system considers users' emoticon selection history to recommend emoticons that each user may like to choose. The procedure of **EBReS-η**, evaluation of this system and the results are described in Sections 3.2, 3.3, and 3.4, respectively.

3.2 Procedure of EBReS-η

EBReS-η basically utilizes two main features: affect analysis system ML-Ask [40] and an originally created emoticon database. Details of ML-Ask and the emoticon database are described in Sections 3.2.1 and 3.2.2, respectively. The system procedure is listed below and the numbers correspond to the numbers in Figure 3.1.

1. A user inputs a sentence that includes an emotive word or phrase.
2. ML-Ask analyzes the input and outputs the emotion type.
3. The emoticon database is reordered based on the emotion type provided by ML-Ask and users' past selection history.

[1] Aspect (1-A): Considering the content of the input when emotive phrase is present and Aspect (2): Considering users' emoticon preference.

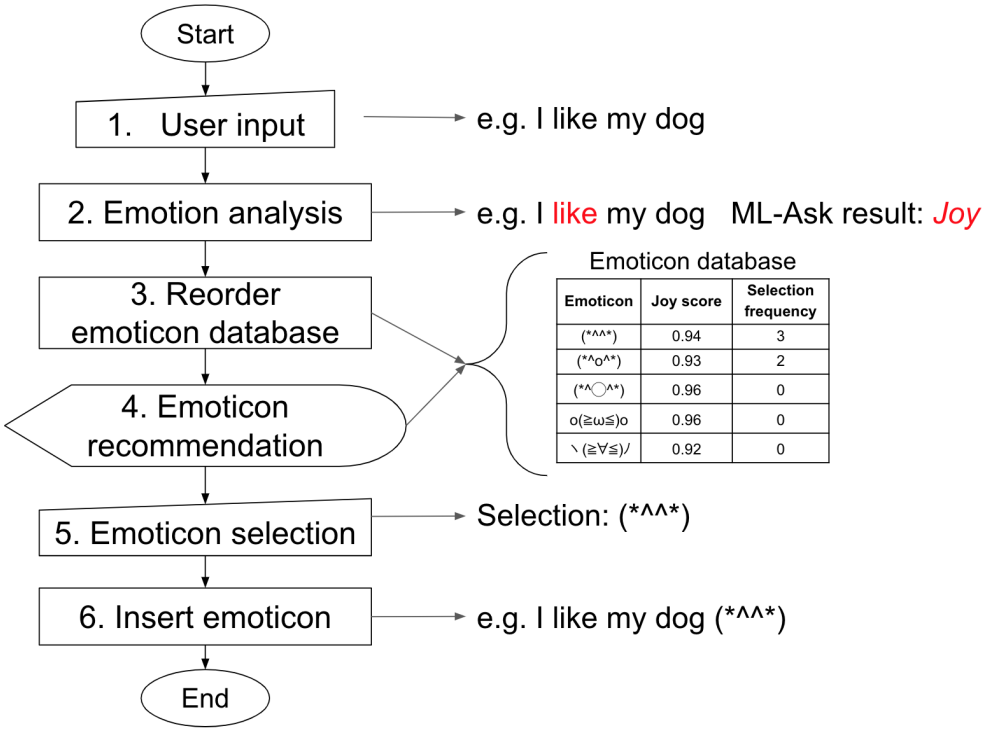


Figure 3.1: System procedure of **EBReS-η**

4. The system shows a list of recommended emoticons to the user.

5. The user selects the desired emoticon from the list.

6. The system automatically inserts the selected emoticon directly after the input message.

To be more concrete, in step 3, the emoticon database first reorders the emoticons based on the number of the selection frequency of each emoticon in descending order and then reorders the emoticons using the scores of the emotion type in descending order. The selection frequency only considers emoticons selected in a sentence that represents the same emotion type as the input sentence. In this way, emoticons that are selected in the past can be recommended at the top of the list, and then the rest of the emoticons are reordered based on the score of the emotion category provided by ML-Ask. The score is the strength of the emotion expressed by the emoticon which is manually annotated. The reason **EBReS-η** does not recommend one emoticon but emoticons in the order shown in Figure 3.1 (3) is because various emoticons express the same emotion category and I believe that users do not always choose the same emoticon but select various emoticons expressing the same emotion type for sentences of similar context.

3.2.1 Affect analysis system: ML-Ask

Ptaszynski *et al.* [40] created an affect analysis system, ML-Ask, for Japanese utterances. ML-Ask firstly separates emotive words from non-emotive words and searches expressions of specific emotion types in the emotive words. The expressions are based on a Japanese emotive word dictionary created by Nakamura [41], in which he categorized emotions into ten different types: JOY[2], LIKE, RELIEF, ANGER, DISLIKE,

[2]Small capital font indicates labels (e.g. emotion type, polarity type, etc.).

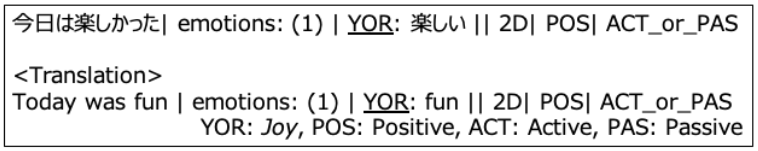


Figure 3.2: Example of the output given by ML-Ask [4]

FEAR, SADNESS, AMAZEMENT, EXCITEMENT, and SHYNESS. ML-Ask was employed in my proposed system because the emoticon database also applied Nakamura's definition of ten emotion types and is completely compatible with ML-Ask.

Figure 3.2 is an example output given by ML-Ask. It generates output as follows: input, number of emotions, emotion type, and emotive word(s) in the input sentence, two dimensions designating if the input expresses as positive or negative, and if the input is active or passive. Only "emotion type" (the underlined part in Figure 3.2) was used from the result to reorder the emoticon database. If there are more than two emotive words and phrases in the input sentence, only the last expression that appeared in the sentence is used because the words that appear in the latter have more influence on selecting emoticons [31].

3.2.2 Emoticon database

The emoticon database was constructed manually utilizing the idea of Kawakami [42]. Kawakami constructed a database of emoticons with emotion and emotive strength to utilize this set for psychological investigation about what type of emotion and its strength can be conveyed just from the appearance of emoticons. The database was created by conducting a questionnaire among 60 Japanese university students. The purpose of the questionnaire was to investigate the power of each of the ten emotions expressed by an emoticon to acquire human perceptions towards this emoticon. The participants rated 59 emoticons by grading the extent to which a given emoticon expressed each of the ten emotion types on a 5-point scale. These emoticons were selected from 140 emoticons registered in iOS 5.0. The emotion types, namely: JOY, LIKE, RELIEF, ANGER, DISLIKE, FEAR, SADNESS, AMAZEMENT, EXCITEMENT, and SHYNESS were defined by Nakamura [41]. These types were classified as polarities, POSITIVE for JOY, LIKE, RELIEF, NEGATIVE for ANGER, DISLIKE, FEAR, SADNESS, and OTHER for AMAZEMENT, EXCITEMENT, and SHYNESS. Finally, the average values of the participants' ratings were calculated, and linked the values to each emoticon to create the database. The condition for the chosen emoticons (i.e. 59 emoticons) was selecting those with no symbols outside of brackets (i.e. "(^0^)" is qualified, but "\(^0^)/" is excluded). In these emoticons, "^" and "0" represent the eyes and mouth of a face, respectively, while the two slashes outside the brackets symbolize arms. Only emoticons that express facial expressions were added to the database to discover what kind of emoticons users choose based on the differences in the facial expressions. The detailed emoticon database is shown in Appendix A.1.

The reason why I chose to conduct a questionnaire to create an emoticon database rather than to collect data automatically from the Internet is that emoticons are mostly used to help people to interpret nuances of a sentence and the level of emotion is not captured by linguistic elements alone [18]. The emoticon analysis system CAO [39] is capable of analyzing the emotion type of the emoticon but does not calculate how strongly the emoticon expresses the emotion which is important in ordering emoticons

Figure 3.3: Example of emoticon ratings in each of the 10 emotions [5]

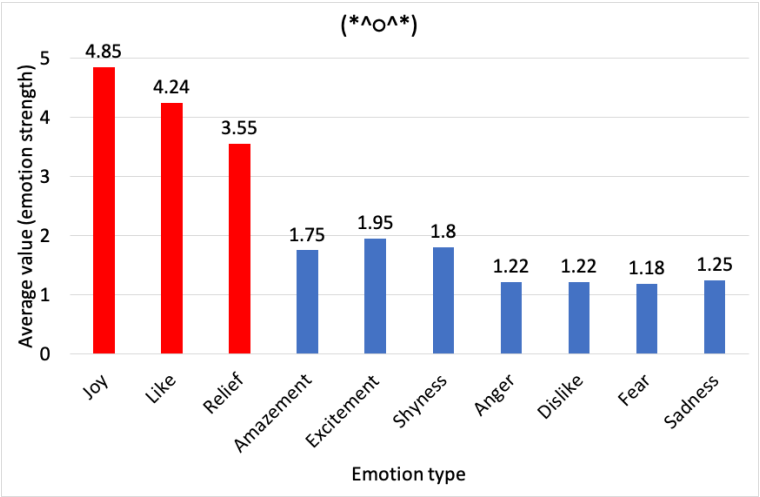


Figure 3.4: Example of an emoticon connected with average values of each of the ten emotions. The red bars are values larger than 3.00 and the blue bars are values smaller than 3.00.

for the recommendation.

Survey

60 university students participated in the questionnaire in July 2012. The participants consisted of 22 men and 38 women, and the average was 20.7 years old. The participants rated 59 emoticons by grading the extent to which a given emoticon expressed each of the ten emotion types on a 5-point scale. The question was phrased as follows: "On a 5-point scale, please rate how well the emoticon expresses each of the ten emotions (1. Not expressed 2. Poorly expressed 3. Neutral 4. Somewhat expressed 5. Expressed)." The example of the rating and indicated words that embody each emotion were demonstrated to the participants.

Survey results

Figure 3.3 illustrates how strongly an emoticon "(ë_ë)" expresses each of ten emotions (Figure 3.3; minimum is 1.00 and maximum is 5.00). The average rating values of each emotion for the 59 emoticons were calculated by collecting the participants' ratings and those values were used to create a database. For **EBReS-η**, this database was used to reorder emoticons in the order of the emotion type in the input message recognized by ML-Ask, putting emoticons that express the larger value of the emotion type at

Table 3.1: Correlation coefficient of ten emotion types calculated from the average ratings of each emoticon. ** indicates the P value lower than 0.05 and * indicates the P value lower than 0.01. P value lower than 0.05 means that the correlation is statistically significant.

Emoticon	1	2	3	4	5	6	7	8	9	10
1. JOY	-	0.989^{**}	0.973^{**}	-0.540^{**}	-0.900	-0.825^{**}	-0.641^{*}	-0.515^{**}	-0.694^{*}	-0.615^{*}
2. LIKE	0.989^{**}	-	0.968^{*}	-0.528^{**}	-0.893	-0.830^{*}	-0.641	-0.527	-0.701	-0.595
3. RELIEF	0.973^{**}	0.968^{*}	-	-0.548^{*}	-0.884	-0.842^{*}	-0.643	-0.542	-0.754	-0.640
4. ANGER	-0.540^{**}	-0.528^{**}	-0.548^{*}	-	0.686^{**}	0.239	0.169	0.034	0.560^{**}	0.234^{*}
5. DISLIKE	-0.900	-0.893	-0.884	0.686^{**}	-	0.796^{**}	0.645	0.351^{*}	0.700	0.643
6. FEAR	-0.825^{**}	-0.830^{*}	-0.842^{*}	0.239	0.796^{**}	-	0.867^{**}	0.463	0.594^{**}	0.651^{**}
7. SADNESS	-0.641^{*}	-0.641	-0.643	0.169	0.645	0.867^{**}	-	0.059	0.256	0.391
8. AMAZEMENT	-0.515^{**}	-0.527	-0.542	0.034	0.351^{*}	0.463	0.059	-	0.623^{*}	0.459
9. EXCITEMENT	-0.694^{*}	-0.701	-0.754	0.560^{**}	0.700	0.594^{**}	0.256	0.623^{*}	-	0.728
10. SHYNESS	-0.615^{*}	-0.595	-0.640	0.234^{*}	0.643	0.651^{**}	0.391	0.459	0.728	-

the top.

From the participants' ratings, 35 out of 59 emoticons scored more than 3.00 in at least two emotion types, which means that those emoticons indicate multiple emotion types strongly. For example, the emoticon "(^◯^)" scored 4.85 points in JOY, 4.24 points in LIKE, and 3.55 points in RELIEF (Figure 3.4, red bar) while the other seven emotion types were below 2.00 points.

The correlation coefficient of the emotion types was computed to determine the closeness of emotion types using the 59 emoticons (Table 3.1). As shown in Table 3.1, three emotion types, JOY, LIKE, and RELIEF scored extremely high positive correlations with each other (JOY & LIKE: $r = 0.989$, JOY & RELIEF: $r = 0.973$, and LIKE & RELIEF: $r = 0.968$). In contrast, three emotion types, DISLIKE, FEAR, and SADNESS scored high positive correlations (DISLIKE & SADNESS: $r = 0.645$, DISLIKE & FEAR: $r = 0.796$, and SADNESS & FEAR: $r = 0.867$) with each other and negative correlations with JOY, LIKE, and RELIEF. Therefore, it can be determined that three emotion types (JOY, LIKE, and RELIEF) and three emotion types (DISLIKE, FEAR, and SADNESS) are polarized from each other. ANGER showed a higher positive correlation with DISLIKE and a negative correlation with JOY, LIKE, and RELIEF. Although ANGER is thought to be a negative emotion, it can be positioned somewhat differently to FEAR and SADNESS. EXCITEMENT showed a positive correlation with DISLIKE, SHYNESS, and AMAZEMENT and negative correlations with three POSITIVE emotion types. EXCITEMENT appears to lie close to NEGATIVE emotion types, however, I believe that this is due to the unbalanced number of emoticons expressing each emotion type in the database (Figure 3.5). The numbers of emoticons that express POSITIVE facial expressions (JOY, LIKE, and RELIEF) were significantly higher than NEGATIVE (SADNESS, FEAR, DISLIKE, and ANGER) or OTHER (EXCITEMENT, AMAZEMENT, and SHYNESS) ones (See Figure 3.5).

3.3 Evaluation experiment

The evaluation experiment was performed to evaluate the performance of **EBReS-η** by comparing the system to other methods. I also evaluated whether Aspects (1) and (2) are important for the system to recommend more appropriate and desired emoticons to users. To investigate the importance of Aspect (1-A) (Considering the content of the input) and Aspect (2) (Considering users' emoticon preference), I

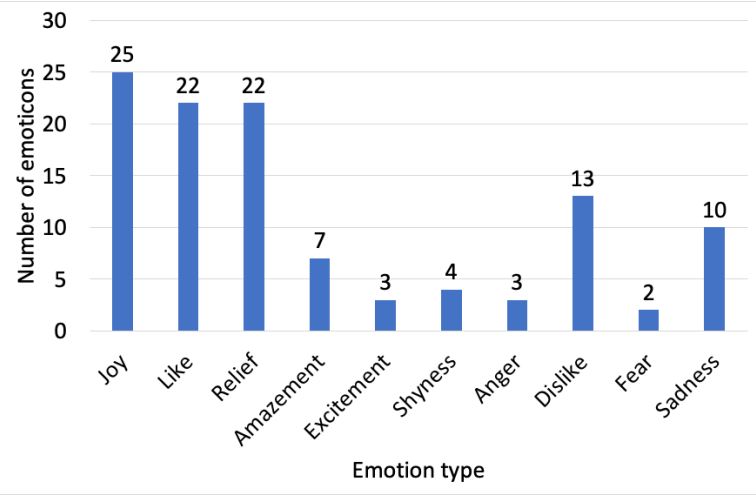


Figure 3.5: Number of emoticons rated more than 3.0 for each emotion type

used the following measurements: (A) average rank of the emoticons chosen by users; (B) proportion of the emoticons chosen by users that were recommended within the top five; (C) user impressions of the system; and (D) user vote for system's performance and for the system that users are fond of. Details of these measurements are described in 3.3.4. Finally, error analysis was performed to investigate for what kind of sentences my system fails to recommend the most preferable emoticons.

The other methods used for comparison were (1) **HiBReS** (**History-Based Emoticon Recommendation System**) and (2) **EBReS** (**Emotion-Based Emoticon Recommendation System**) [5]. (1) **HiBReS** is a method identical to the one currently used in the iOS and Android Japanese keypad. The name "**HiBReS**" is just used in this paper to shorten the system name. This system recommends emoticons to users based on their past emoticon selections and does not regard the content of input messages. (2) **EBReS** recommends emoticons by reordering the database based on the emotion types analyzed from input messages. **HiBReS** and **EBReS** methods were used for comparison because my proposed method is a hybrid of these two methods and it can be simply investigated how the combination of these two methods perform differently from these methods utilizing one feature (user selection history or emotive score of the input emotion type) for emoticon recommendation.

I developed these three methods as a single smartphone application to conduct an evaluation experiment. The application has an interface showing three buttons, each of them linked to one of the three methods (details are introduced in 3.3.2). Participants were asked to use all three methods based on the procedure described in 3.3.3. The application is usable on the iPhone (from iOS 7.0 to iOS 7.0.4). The device used for the experiment was iPhone 5S (iOS 7.0.4). The experiment was performed over nine days from October 31st to November 8th in 2013 with the cooperation of 30 Japanese undergraduate and master course students.

3.3.1 Participants

30 undergraduates and master course students participated in the experiment. The group consisted of 15 men and 15 women (Table 3.2). Their average age was 22.4 years old (standard deviation = 1.77). Among the 30 participants, 60.0% of the students possessed an iPhone or iPad, 33.3% possessed an Android device, and the rest possessed feature phones. Moreover, 86.7% of the students reported that

Table 3.2: Details of participants regarding their device types used to communicate and their daily message and emoticon usage

Measure		Men	Women	Total
Device used daily	iPhone/iPad	6 (20.0%)	12 (40.0%)	18 (60.0%)
	Android	7 (23.3%)	3 (10.0%)	10 (33.3%)
	Others	2 (6.7%)	0 (0.0%)	2 (6.7%)
Daily message usage	Very often	5 (16.7%)	11 (36.7%)	16 (53.3%)
	Somewhat	7 (23.3%)	3 (10.0%)	10 (33.3%)
	Hardly	3 (10.0%)	1 (3.3%)	4 (13.3%)
	Not at all	0 (0.0%)	0 (0.0%)	0 (0.0%)
Daily emoticon usage	Very often	4 (13.3%)	9 (30.0%)	13 (43.3%)
	Somewhat	8 (26.7%)	6 (20.0%)	14 (46.7%)
	Hardly	3 (10.0%)	0 (0.0%)	3 (10.0%)
	Not at all	0 (0.0%)	0 (0.0%)	0 (0.0%)

they "very often" or "somewhat often" send messages daily, and 90.0% use emoticons "very often" or "somewhat often" when sending text messages (e.g. e-mail, SMS, etc.).

3.3.2 Emoticon recommendation application

The emoticon recommendation application performs as follows:

1. First, a user selects which method to use (Figure 3.6, 1).
2. When the user touches the textbox (Figure 3.6, 2, red frame), a Japanese keypad appears (Figure 3.6, 3, red frame).
3. The user inputs a sentence (Figure 3.6, 4, red frame). For example in Figure 3.6, 4, the input message is "今日は楽しかった" (*kyo ha tanoshikatta*, meaning "I had a fun day today").
4. When the user touches the "顔文字" (*kaomoji*, meaning "emoticon") button (Figure 3.6, 5, red frame), a list of emoticons appears (Figure 3.6, 6, red frame); The order of the emoticons is decided using each method (**EBReS-η**, **HiBReS**, and **EBReS**).
5. The user selects an emoticon by scrolling up and down the list, and touches the "決定" (*kettei*, meaning "done") button to insert it (Figure 3.6, 6, green frame).
6. The system automatically inserts the chosen emoticon right after the input (Figure 3.6, 7, red frame).

3.3.3 Experiment procedure

The experiment was conducted as follows:

1. Respondents firstly filled out basic information: their university year group, gender, age, faculty, whether they possess a smartphone, and whether they use emoticons daily when sending messages (via e-mail, SMS, etc.).

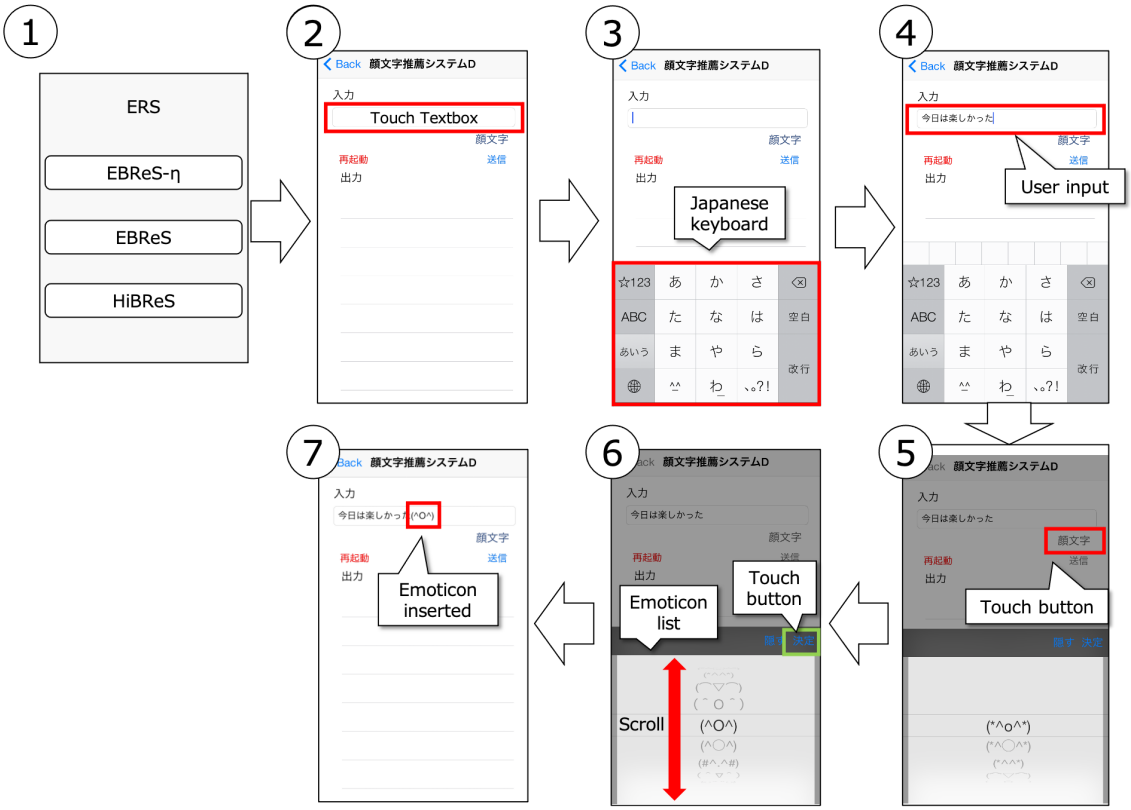


Figure 3.6: Emoticon recommendation procedure in the application

2. Respondents tested one of the three methods; the order in which a participant tested these methods was decided randomly.

3. Respondents rated the method using 22 impression word pairs on a 7-point scale; the 22 impression word pairs are described in 3.3.4.

4. Respondents tested the remaining methods as described above in steps 2 and 3.

The input messages were prepared beforehand (a list of 15 sentences with one emotive word) by the author based on the following procedure. The sentences were chosen from 300 input sentences made by participants' during the previous experiment explained in [5]. These sentences were examined beforehand by asking 10 Japanese students to rate them on a 5-point scale (1.0 point to 5.0 point) and chose sentences that express one of the 10 emotion types higher than 4.0 point. Examples are shown in Table 3.3, and all the sentences are shown in Appendix A.2. The average points for the example sentences are shown in brackets next to the emotion types in Table 3.3. The list of 15 sentences was comprised of five sentences with POSITIVE emotions (one of the emotions from JOY, LIKE, and RELIEF), five NEGATIVE emotions (one of the emotions from ANGER, DISLIKE, FEAR, and SADNESS), and five OTHER emotions (one of the emotions from AMAZEMENT, EXCITEMENT, and SHYNESS). The emotions used in the list were chosen randomly from polarities, POSITIVE, NEGATIVE, and OTHER. The participants were asked to enter each sentence in every three methods.

Table 3.3: Example of sentences shown to participants. Emotive expressions are highlighted in bold.

Japanese sentence	Translation	Emotion type
その漫画は**好き**だよ	I **like** this comic book	LIKE (4.9)
それはちょっと**恥ずかしい**	This is a little **embarrassing**	SHYNESS (4.3)
怯えてしまう	I am **frightened**	FEAR (4.8)

3.3.4 Evaluation measurements

(A) Average rank of the chosen emoticons

The rank of the chosen emoticons was examined to find out which of these methods can recommend emoticons more at the top of the list. The equation to calculate the average rank is shown in Equation 3.1. In this equation, N denotes the number of recommendations given by a system and r_i is the rank of the i-th recommendation made by the system. By using Equation 3.1, a lower value means that the system can recommend more emoticons at the top of the list (minimum is 0 and maximum is 59).

$$AverageRank = \frac{\sum_{i=1}^{N} r_i}{N} \tag{3.1}$$

(B) Proportion of chosen emoticons appeared within the top five

The proportion of chosen emoticons that appeared within the top five was used as one of the evaluation measurements to discover how many of the emoticons were actually selected among the top five. The equation to calculate this proportion for each method is shown in Equation 3.2, whose N is the number of recommendations given by a system and r_i is the rank of the i-th recommendation. If the r_i is lower than 5, 1 is added, and 0 for otherwise, simply counting how many times the user chose an emoticon that was displayed within the top five recommendations. The higher value means that more emoticons chosen by users were displayed within the top five recommendations (minimum is 0% and the maximum is 100%).

$$Proportion = \frac{\sum_{i=1}^{N} f(r_i)}{N} \times 100, f(r_i) = \begin{cases} 1, & \text{if } r_i \leqq 5. \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

(C) Users' impression towards the systems

Semantic differential (SD) scale was adopted to examine users' intuitions towards the systems under evaluation. SD scale was created by Osgood *et al.* [43] and is used to investigate users' impressions of an object (e.g. a system, place, etc.). In this method, the subject's impression of each system is quantified on a 7-point scale using various bipolar words (e.g. "bad" and "good"). In our evaluation, 22 impression word pairs (see Table 3.4) suitable to evaluate a computer system were chosen from a survey of what kind of word pairs are used in various research domains, conducted by Inoue *et al.* [44]. In this evaluation, I examined the scores for each impression word pairs collected from the participants for each system and also performed factor analysis [43] to map users' impression onto a smaller number of factors to visualize their impressions towards each system. The details of factor analysis are described in Appendix A.3.

Table 3.4: 22 impression word pairs

Aspect	Abbreviation	Left word	Right word
Amusement	Amu	Boring	Fun
Impressiveness	Imp	Unimpressive	Impressive
Friendliness	Fri	Unfriendly	Friendly
Handiness	Han	Difficult to use	Easy to use
Speed	Sp	Slow	Fast
Convenience	Con	Inconvenient	Convenient
Necessity	Nec	Unnecessary	Necessary
Lightness	Lig	Heavy	Light
Recognizability	Rec	Obscure	Clear
Fondness	Fon	Dislike	Like
Novelty	Nov	Old	New
Simplicity	Sim	Complicated	Simple
Interest	Int	Uninterested	Interested
Freshness	Fre	Mediocre	Novel
Accurate	Acc	Inaccurate	Accurate
Useful	Use	Useless	Useful
Visibility	Vis	Difficult to see	Easy to see
Easiness	Eas	Difficult	Easy
Choosability	Cho	Difficult to choose	Easy to choose
Specialness	Spe	Ordinary	Special
Smartness	Sma	Dumb	Smart
Satisfactory	Sat	Unsatisfied	Satisfied

先ほど使用していただいた３つのシステムに対し、次の条件で順位をつけてください。
(Please rank the systems based on the following questions)

- どのシステムが優れているか (Rank the system in the order of their performance)

1位 (1st) EBReS-η	2位 (2nd) EBReS	3位 (3rd) HiBReS

- どのシステムが好きか (Rank the system in the order of your preference)

1位 (1st) EBReS-η	2位 (2nd) EBReS	3位 (3rd) HiBReS

Figure 3.7: Example of a vote for system's performance and users' preferences

(D) User vote for systems' performance and preference

Users' votes on the system's performance and their preferences were collected after participants finished the experiment. An example of a vote is shown in Figure 3.7. As shown in this figure, I asked the participants to write down the names of the systems in the order of their performance and users' preferences. The system name which is written next to "1 位" (*ichii*, meaning "1st place") is the system in which performance is the best and the system that a user prefers the most. The system name written next to "3 位" (*san'i*, meaning "3rd place") is the system that is evaluated as the worst in relation to performance or preference. I collected the participants' votes and counted which system had acquired the best position scores (from 1 to 3).

3.4 Results and discussion

3.4.1 (A) Average rank & (B) proportion of emoticons chosen within the top five

Figure 3.8 shows the number of chosen emoticons by rank (bars), the proportion of emoticons chosen from the top five candidates (horizontal curly brace), and the average ranking of chosen emoticon for each system (red line). As can be seen in the figure, 72.3% of the emoticons were selected from within the top five when participants tested **EBReS-η** which yields the best results among all systems (**EBReS**: 58.0%, **HiBReS**: 30.0%). Also, the average rank of the chosen emoticons was 5.2, whereas the average ranks were 6.8 and 17.4 for **EBReS** and **HiBReS**, respectively. From these results, it can be concluded that recommending based on the emotion type of the emotional phrase recognized in the input and on user emoticon preference can recommend desired emoticons more often at the top of the list.

3.4.2 (C) Participants impression using SD scale

Figure 3.9 shows the result of participants' impression for each 22-impression pairs introduced in Table 3.4. The dots in the figure show the average impression score calculated from participants' ratings. The dots closer to "1" have strong impressions of the words on the left, whereas the dots closer to "7" are better

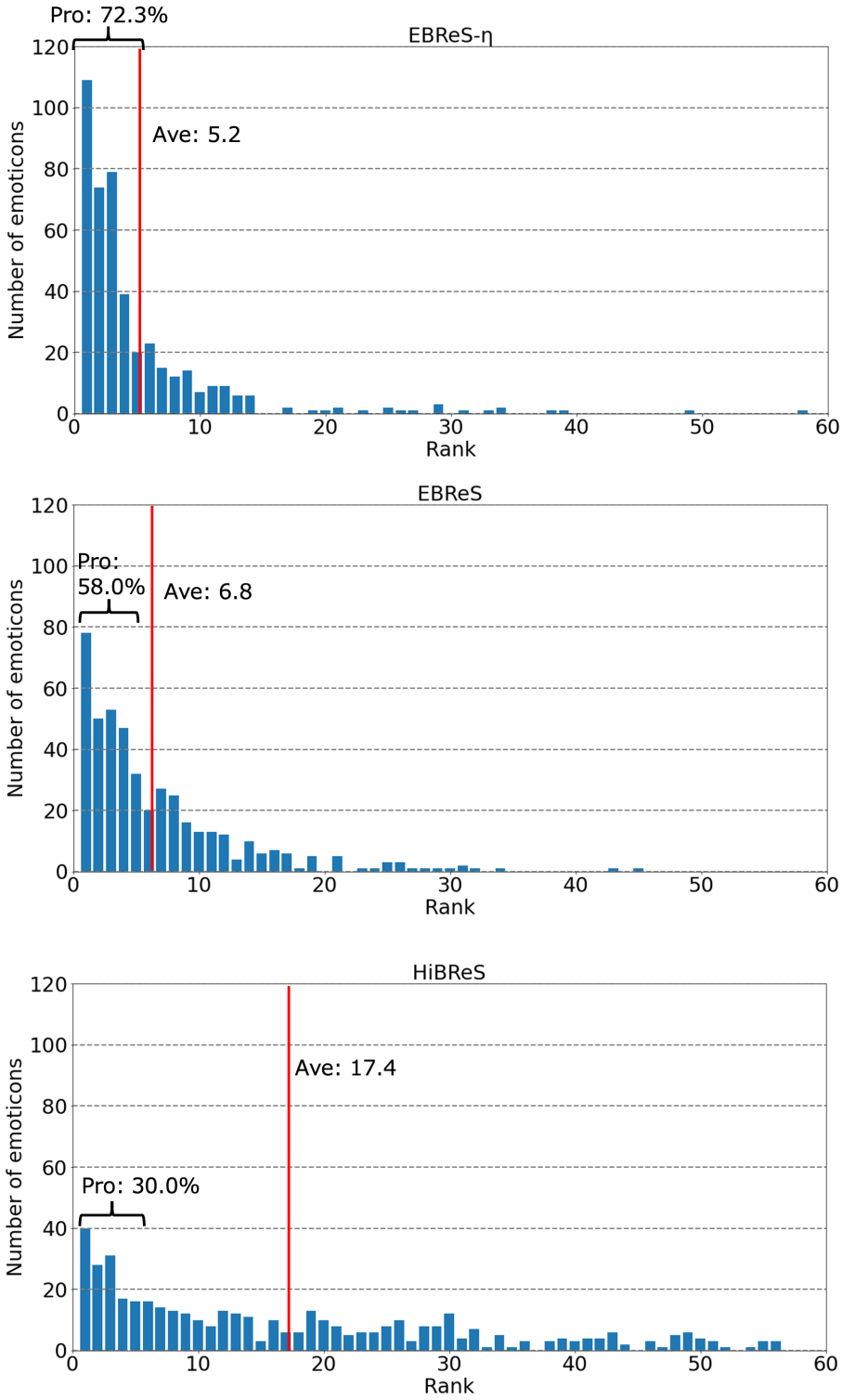


Figure 3.8: Number of chosen emoticons by rank. "Pro" stands for "Proportion of emoticons chosen within the top five". "Ave" stands for "Average ranking of the chosen emoticons". Red line indicates the average ranking.

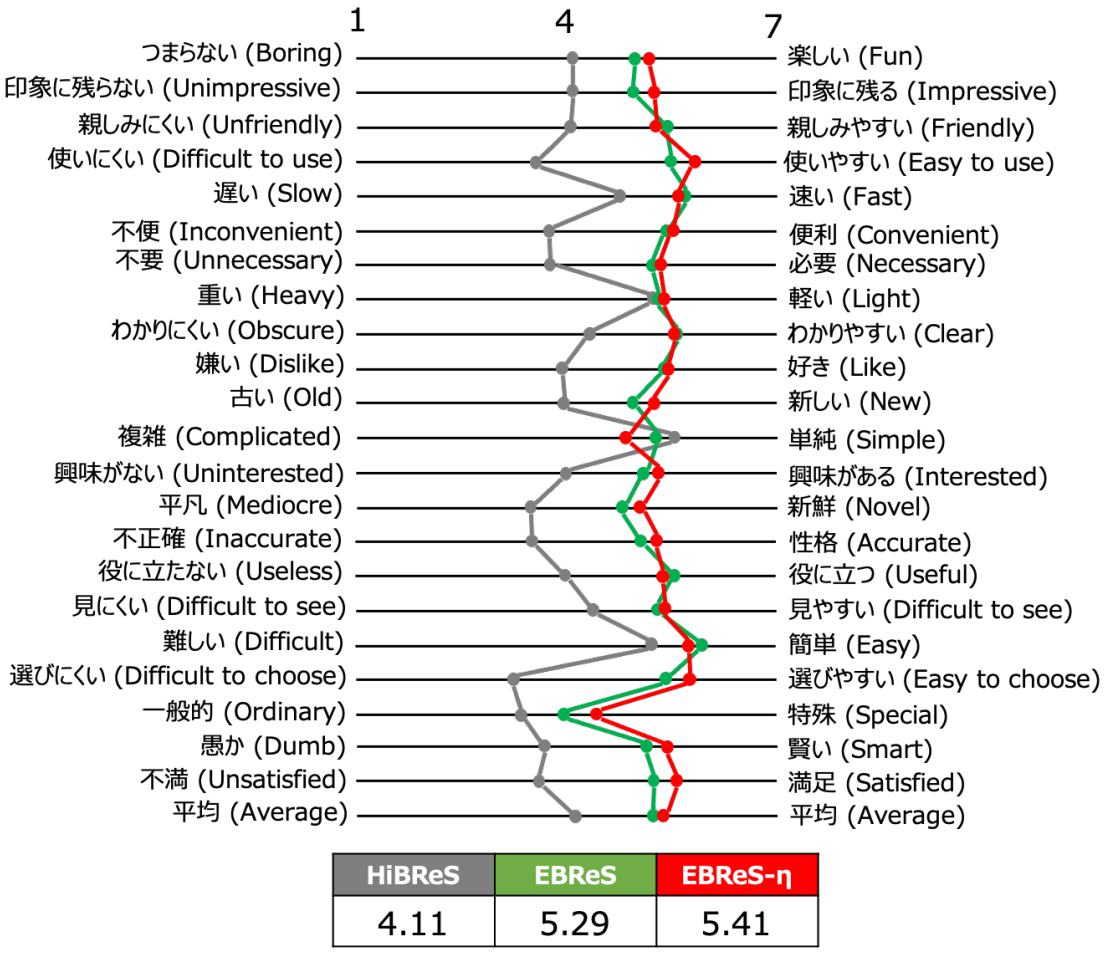


Figure 3.9: Participants impression of each system using 22 impression word pairs

Table 3.5: Factor loadings of each of the 22 impression word pairs in the semantic differential (SD) scale

22 impression word pairs		Factor 1	Factor 2	Factor 3	Factor 4
Left word	Right word	Usefulness	Novelty	Simplicity	Process quality
Difficult to use	Easy to use	**0.991**	-0.097	-0.035	0.005
Difficult to choose	Easy to choose	**0.972**	-0.016	-0.165	0.055
Unsatisfied	Satisfied	**0.971**	0.062	-0.113	-0.004
Inconvenient	Convenient	**0.924**	-0.001	-0.035	0.104
Unnecessary	Necessary	**0.907**	0.071	0.060	0.006
Useless	Useful	**0.891**	0.027	-0.045	0.054
Dislike	Like	**0.842**	-0.005	0.091	-0.007
Inaccurate	Accurate	**0.833**	0.144	-0.190	-0.108
Obscure	Clear	**0.722**	-0.114	0.177	0.109
Unfriendly	Friendly	**0.716**	-0.028	0.212	-0.040
Difficult to see	Easy to see	**0.623**	-0.068	0.163	0.113
Dumb	Smart	**0.616**	0.386	-0.069	0.027
Uninterested	Interested	**0.431**	0.321	0.364	-0.165
Unimpressive	Impressive	**0.388**	0.320	0.337	-0.189
Ordinary	Special	-0.142	**0.951**	-0.169	0.166
Mediocre	Novel	0.014	**0.811**	0.095	0.083
Old	New	0.202	**0.686**	0.053	0.060
Complicated	Simple	-0.248	-0.135	**0.548**	0.127
Boring	Fun	0.356	0.178	**0.518**	-0.140
Difficult	Easy	0.327	-0.024	**0.426**	0.191
Slow	Fast	0.051	0.203	-0.012	**0.814**
Heavy	Light	-0.020	-0.016	0.230	**0.585**

characterized by the words on the right. The average impression score is shown below each system name in the figure. As can be seen in the figure, **EBReS-η** showed the highest scores among the three systems for 17 impression word pairs out of 22 impression word pairs. The average impression score for **EBReS-η** was 5.41 (minimum is 1.0 and maximum is 7.0). The second highest was **EBReS**, which was slightly lower than the **EBReS-η** by 0.12 points on average. From these results, it is clear that recommending emoticons according to emotion types recognized in user input simplifies the process of choosing emoticons. It was also discovered that **EBReS-η** (4.90 points) and **EBReS** (5.40 points) scored lower than **HiBReS** (5.60 points) for the impression word pair "Complicated $\Leftrightarrow$ Simple." I assumed that most participants rated this by considering the simplicity of the **HiBReS** method that this method only recommends emoticons that were selected in the past. Detailed ratings given by the participants are shown in Appendix A.4.

I carried out factor analysis to condense a large number of variables into a few interpretable underlying factors and summarize the respondents' perceptions towards each of the three systems. The factor analysis resulted in four factors with eigenvalues exceeding 1.0 which accounted for 65.0% of the variance. Table 3.5 shows the factor loadings for the 22 impression word pairs. The first factor is made up of 14 scales and

Table 3.6: Rankings on the systems' performance and users' preferences. The highest value for each method is emphasized with bold.

	Performance			Preference		
System	1st	2nd	3rd	1st	2nd	3rd
EBReS-η	**23**	6	1	**21**	8	1
EBReS	6	**16**	8	6	**14**	10
HiBReS	1	8	**21**	3	8	**19**

can be described as "Usefulness" as it contains word pairs such as "Difficult to use ⇔ Easy to use", "Inconvenient ⇔ Convenient", etc. The second factor is made up of three scales ("Ordinary ⇔ Special", "Mediocre ⇔ Novel", and "Old ⇔ New"). These word pairs can be summarized as "Novelty". The third factor was also comprised of three scales ("Complicated ⇔ Simple", "Boring ⇔ Fun", "Difficult ⇔ Easy"), therefore I summarized these scales as "Simplicity". The fourth factor contained two scales ("Slow ⇔ Fast" and "Heavy ⇔ Light)" which indicates about system process quality, therefore "Process quality" was named. The detailed process of factor analysis using R language is shown in Appendix A.4.

Figures 3.10 and 3.11 show factor loadings of each impression word and average factor scores of respondents categorized by each system. As shown in Figure 3.10, **EBReS-η** ("**EBReS-η**" in Figure 3.10) demonstrated the highest impression for "Usefulness" and "Novelty" among the three systems, whereas **HiBReS** ("**HiBReS**" in Figure 3.10) was ranked by far the lowest in both these aspects. **EBReS** produced a lower impression for both these aspects than **EBReS-η** ("**EBReS-η**" in Figure 3.10). As shown in Figure 3.11, **EBReS** and **EBReS-η** demonstrated higher than 0 for both axes and a more positive impression than **HiBReS**. However, the impression towards the proposed method **EBReS-η** was negative than the **EBReS** method. Overall, these results can imply that the impression from respondents towards systems that put more importance on the emotion type retrieved from input messages for recommendation is thought to perform better than just recommending emoticons based on users' selection history.

3.4.3 (D) User vote for system's performance and preference

Respondents were asked to vote on which system performed best and which system they prefer to use. Table 3.6 shows the results of this ranking. As shown in Table 3.6, 23 out of 30 participants ranked **EBReS-η** as performing the best, 16 participants ranked **EBReS** as second, and 21 participants ranked **HiBReS** as the third. The order of the participants' preference toward these systems was **EBReS-η** (21 participants voted this system as first), **EBReS** (14 participants voted this system as second), and **HiBReS** (19 participants voted this system as third). From these results, **EBReS-η** performed the best and preferred by the respondents than the other two systems.

3.4.4 Error analysis

My proposed method recommends emoticons of the same emotion type to the emotive phrase in the input sentence. This approach was based on the observation that users choose emoticons expressing the same polarity or emotion type to the sentence in order to enhance the meaning [15, 20]. In this investigation,

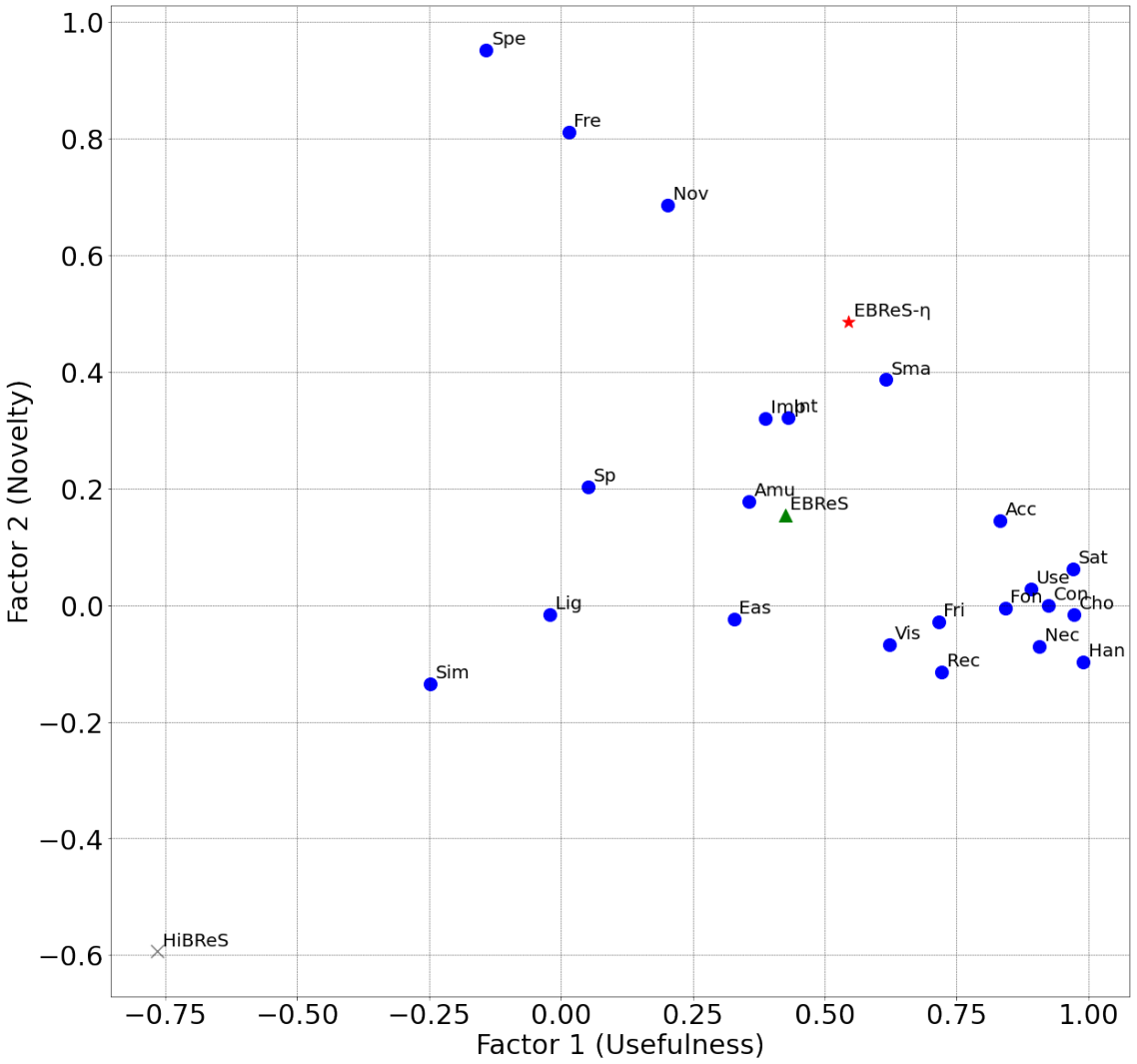


Figure 3.10: Biplot of the two factor models for factor 1 (Usefulness) and 2 (Novelty). Factor 1 explains 42.0% and factor 2 explains 12.0% of the total variance. Red star = **EBReS-η**, green triangle = **EBReS**, and gray cross = **HiBReS**. Blue dots indicate abbreviations for SD adjectives: Amu = Amusement, Imp = Impressiveness, Fri, Friendliness, Han = Handiness, Sp = Speed, Con = Convenience, Nec = Necessity, Lig = Lightness, Rec = Recognizability, Fon = Fondness, Nov = Novelty, Sim = Simplicity, Int = Interest, Fre = Freshness, Acc = Accurate, Use = Useful, Viz = Visibility, Eas = Easiness, Cho = Choosability, Spe = Specialness, Sma = Smartness, Sat = Satisfactory.

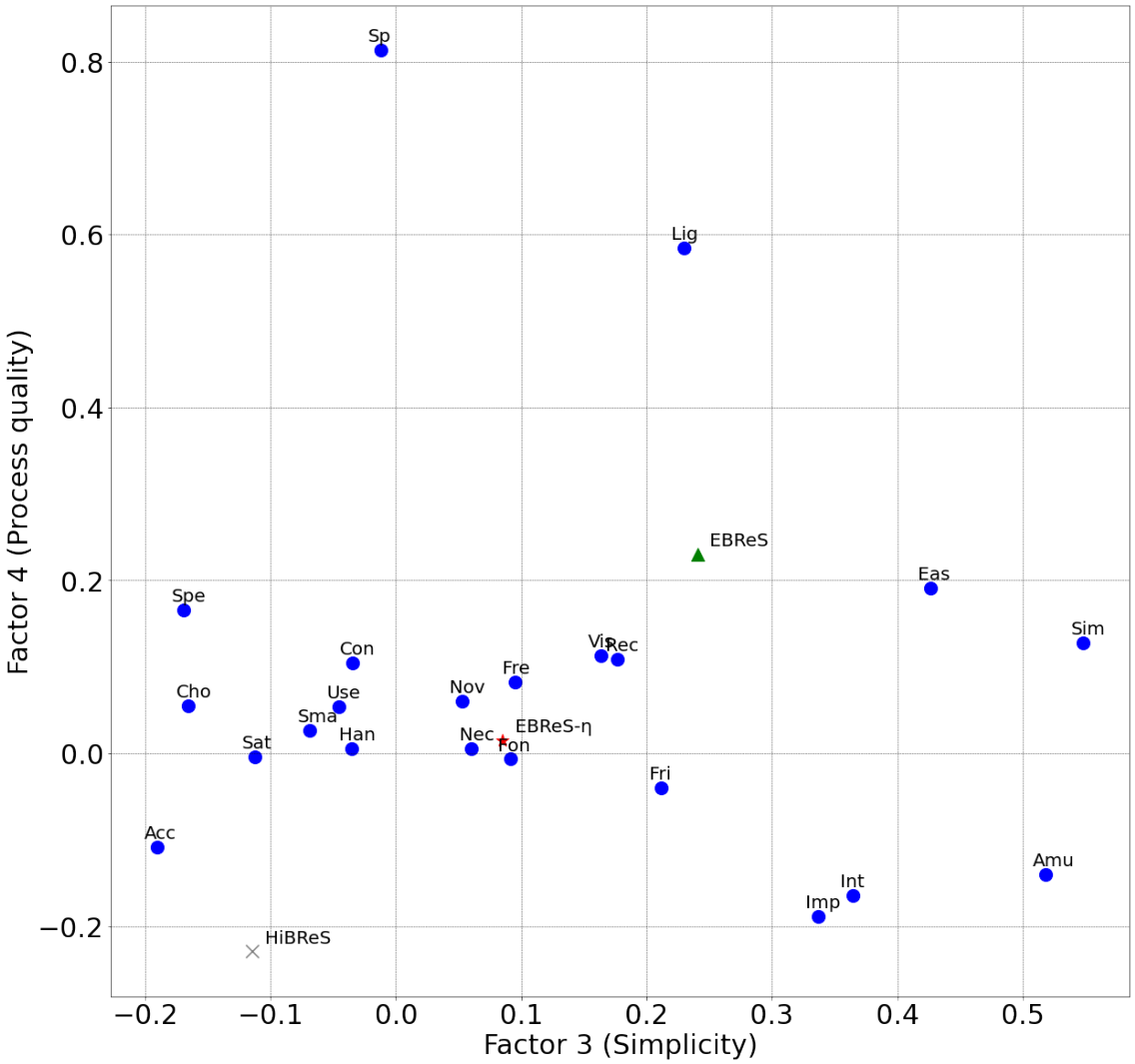


Figure 3.11: Biplot of the two factor models for factor 3 (Simplicity) and 4 (Process quality). Factor 3 explains 5.9% and factor 4 explains 5.6% of the total variance. Red star = **EBReS-η**, green triangle = **EBReS**, and gray cross = **HiBReS**. Blue dots indicate abbreviations for SD adjectives: Amu = Amusement, Imp = Impressiveness, Fri, Friendliness, Han = Handiness, Sp = Speed, Con = Convenience, Nec = Necessity, Lig = Lightness, Rec = Recognizability, Fon = Fondness, Nov = Novelty, Sim = Simplicity, Int = Interest, Fre = Freshness, Acc = Accurate, Use = Useful, Viz = Visibility, Eas = Easiness, Cho = Choosability, Spe = Specialness, Sma = Smartness, Sat = Satisfactory.

Table 3.7: Number of sentences in which the emotion type of selected emoticons and emotive words in sentences disagreed. POSITIVE, NEGATIVE, and OTHER are the polarity types of the emotive phrase recognized within sentences.

Total	POSITIVE	NEGATIVE	OTHER
320	39	43	238

Table 3.8: Examples of sentences in which the polarity types of emotive phrase and the selected emoticon disagreed with each other. Emotive phrase in the sentence is emphasized with bold.

Sentence	Translation	Emoticon	Polarity type	
			Sentence	Emoticon
起きたら 8 時で**焦った**	I had to **rush** because it was 8 am when I woke up	(; ´ 凵 `)	OTHER	NEGATIVE
試験が終わったし**気楽だ**	Exams are over, I'm so **relaxed**	(T _ T)	POSITIVE	NEGATIVE
昨日の夢は**怖かった**	My yesterday dream **scared** me	(; ￣ O ￣)	NEGATIVE	OTHER

I analyzed the disagreement between the polarity type of the chosen emoticon and the emotive phrase that appeared in the input sentence in order to investigate for what kind of sentences users tend to choose emoticons of the different polarity type. Among 1,350 sentences (15 sentence × 3 method × 30 participant), emoticons selected for 320 sentences (23.7%) expressed polarity type different from the emotive phrase contained in the input sentences (Table 3.7). Moreover, 238 sentences contained words categorized as OTHER (AMAZEMENT, EXCITEMENT, and SHYNESS). Table 3.8 shows example sentences in which the polarity types between the emotive phrases and the selected emoticons disagreed with each other. For example, the sentence, "起きたら 8 時で焦った" (*okitara 8jide asetta*, meaning "I had to rush because it was 8 am when I woke up") contains an emotive phrase "焦る" (*aseru*, meaning "rush" or "be in a hurry") which is categorized as OTHER (EXCITEMENT). However, the NEGATIVE (DISLIKE) emoticon "(; ´ 凵 `)" was selected for this sentence. I assume that the user wanted to express the negative feeling clearly with the word "焦る" (*aseru*, meaning "rush" or "be in a hurry"), therefore, the NEGATIVE emoticon was selected. The second example sentence, "試験が終わったし気楽だ" (*shikenga owattashi kirakuda*, meaning "Exams are over, I'm so relaxed") was inserted with an emoticon "(T_T)". I assume that the user wanted to express the overwhelming feeling with the "burst into tears" emoticon rather than selecting a happy emoticon. For the third example, " (; ￣ O ￣) " was chosen for the sentence, "昨日の夢は怖かった" (*kinouno yumeha kowakatta*, meaning "My yesterday dream scared me"). I assume that the user wanted to express the feeling of being surprised instead of expressing fear.

3.4.5 Discussion

Overall results show that combining elements of "recommending based on the emotive content of the input message" and "recommending based on the users' selection history" are effective to suggest more desired emoticons at the top of the list than only using one of them for a recommendation. However, there were cases when users selected emoticons that express different emotions from an emotive word recognized in the input sentence (see 3.4.4 for details). Therefore, I believe that it is not sufficient to just recommend emoticons of the same emotion type for the input messages and it is better to recommend emoticons by learning how users select them for certain input messages. In order to obtain better recommendation

results, I decided to apply machine learning techniques to discover what kind of emoticons are tend to be used in certain sentences and recommend emoticons depending on such data. This part of my research is introduced in Chapter **4**.

3.5 Conclusion of Chapter 3

In this chapter, I presented an emoticon recommendation method, **EBReS-η** that recommends emoticons based on emotive statements contained in input messages and on the users' past emoticon selection. This method analyzes emotions from user-entered sentences by using the affect analysis system ML-Ask, reorders the emoticon database based on the emotive score of the emotion type from the analytical result and on the previous selections made by the user. The emoticon database was also created by utilizing an idea proposed by Kawakami [42], and comprised of 59 emoticons with the points expressed from each of ten distinctive emotions.

Evaluation experiments were performed to compare the performance of **EBReS-η** to other two systems: **HiBReS** (**History-Based emoticon Recommendation System**) and **EBReS** (**Emotion-Based emoticon Recommendation System**). Conducted experiments demonstrated that **EBReS-η** performed best showing that 72.3% of emoticons were chosen among the top five recommendations by the respondents. The results for the other methods were, 58.0% for **EBReS** and 30.0% for **HiBReS**. It was also discovered that the average rankings of the chosen emoticon scored higher in the following order: 5.2 (**EBReS-η**), 6.8 (**EBReS**), and 17.4 (**HiBReS**). These quantitative results strongly suggest that recommending emoticons based on the emotion type of emotive statements in the input messages play a great role in recommending desired emoticons to users. Moreover, recommending users' frequently chosen emoticons considering the emotion type in the input performs better than just recommending based on the emotion type. The average user impression towards the proposed system was most positive for "Usefulness" and "Novelty", and less positive than the **EBReS** method for "Simplicity" and "Process performance". The proposed method also achieved the best results in voting, showing that 23 out of 30 respondents declared **EBReS-η** as the best performing algorithm and 21 respondents answered that they wish to use **EBReS-η** most among all three systems.

This system focused on recommending suitable emoticons based on the emotive statements in input messages, however, there are messages in which users do not insert emoticons with the same emotion type. Also, some text messages in the real world do not contain clear emotive expressions. Therefore, in Chapter **4**, I propose a method that learns how emoticons are used in sentences and is also capable of recommending emoticons when emotive phrases do not exist in the input (Aspect (1-B)).

Chapter 4

Deep Emoticon Recommendation System (Deep-ERS)

4.1 Introduction

This chapter describes a proposed system **Deep-ERS** (**Deep-Emoticon Recommendation System**) that recommends suitable emoticons by analyzing the whole input sentence. This is different from my previous method (see Chapter **3** for details) which only focused on an emotive phrase recognized in input messages. This system is also capable of recommending desired emoticons even when a message does not contain any emotive statement (Aspect (1-B)). The outline of Chapter **4** is as follows: in the next section, I will explain the motivation of this system based on the analytical result of emoticon usage in sentences collected from the Internet. 4.3 describes the procedure of **Deep-ERS** and the pre-trained models that were employed in this system. 4.4 explains an evaluation experiment to measure the performance of the **Deep-ERS** compared with conventional methods. Results and discussion are introduced in 4.5.

4.2 Preliminary investigation

The purpose of investigating emoticon usage in Japanese sentences is to learn the limitations of the rule-based methods (described in Chapter **3** and [30, 31]) which use emotive cues in sentences for emoticon recommendation. I performed two analyses: (1) the percentage of inputs that did not contain emotive phrases and (2) the percentage of data where the polarity/emotion type of emotive phrases did not match with the inserted emoticons. These analyses were performed to discover the number of input sentences in the real world which cannot recommend emoticons from the rule-based methods.

4.2.1 Data description

652,924 pairs of sentences and emoticons were collected from Twitter[1], Nihongo Web Corpus (NWC) [45], and YACIS [46]. Examples are shown in Table 4.1. These corpora were selected because emoticons occur frequently in casual messages on social media and in blogs. The Twitter data was collected between

[1] `https://twitter.com`

Table 4.1: Examples of sentence-emoticon pairs. Emotive words are emphasized in bold.

Sentence	Emoticon	Sentence		Emoticon	
		polarity	emotion	polarity	emotion
僕もイベントが**好き**です (I also **like** events)	(^_^)	POSITIVE	LIKE	POSITIVE	JOY
もう一回順番待ちするの**面倒**くさいなあ (It is **troublesome** to wait again)	(;_;)	NEGATIVE	DISLIKE	NEGATIVE	SADNESS
野宿できるとこ探さないと (I have to look for a place to sleep)	(ーー；)	NONE	NONE	NEGATIVE	DISLIKE

March 16, 2010, and January 10, 2011. NWC is text data collected via Yahoo! Web search API[2] between June and September 2010. YACIS data was collected from Ameba blog service[3] in December, 2009. Some data contained multiple sentences merged into one line, therefore, these sentences were divided into each sentence by periods, exclamation marks, question marks, emoticons, and emojis. Only data that contained at least five tokens (POS chunks) of natural language text followed by one emoticon was used in the investigation. Sentences that contained less than five tokens were not collected because these were too short and therefore difficult to have any emoticons suggested. The average number of tokens in one sentence was 10.3 and the longest sentence contained 261 tokens.

Annotation of polarity and emotion types to sentences

Polarity and emotion types were automatically annotated to each sentence using ML-Ask [40] (see 3.2.1 for details). The polarity in ML-Ask is defined as POSITIVE, NEGATIVE, and NEUTRAL categories. Emotion types are grouped as follows: JOY, LIKE, and RELIEF belong to POSITIVE, ANGER, DISLIKE, FEAR, and SADNESS to NEGATIVE, and AMAZEMENT, EXCITEMENT, and SHYNESS to NEUTRAL categories. However, words expressing AMAZEMENT, EXCITEMENT, and SHYNESS can be either positive or negative depending on the context, therefore, these data were annotated as OTHER instead of NEUTRAL. When there were no emotive phrases in a sentence, polarity and emotion categories were marked as NONE. If there are multiple emotive phrases in a given input, the polarity and emotion type of the last emotive word or phrase that appeared in the sentence was used because the last appeared ones tend to influence the selection of the emoticon most in the Japanese language [31].

Emotion database

The emoticon database [8] used in this analysis contained 299 different emoticons including ones that were used in Chapter **3**. Details of the emoticon set are shown in Appendix B.1. 59 emoticons also used in Chapter **3** were the ones that do not contain any symbols outside of the brackets and the rest of the 240 emoticons contained symbols outside of the bracket. 240 emoticons were collected from "顔文字カフェ" [47] (*kaomojikafe*, meaning "Emoticon Cafe"), a website that introduces many different kinds of emoticons. Initially, 839 emoticons were collected from "シンプル顔文字" (*shimpuru kaomoji*, meaning "simple emoticons") section and three graduate school students (two male and one female students) were

[2]Yahoo! Web search API: `https://developer.yahoo.co.jp/changelog/2013-08-14-search.html`
[3]Ameba blog: `https://ameblo.jp/`

Table 4.2: Example emoticons from emoticon dictionary introduced in [8]. The highest value for each emoticon is emphasized in bold.

Emoticon	JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SADNESS	AMAZEMENT	EXCITEMENT	SHYNESS
(∗^_^∗)	**0.87**	0.78	0.65	0.10	0.08	0.06	0.09	0.10	0.16	0.25
(T_T)	0.08	0.08	0.12	0.20	0.51	0.47	**0.87**	0.19	0.25	0.29
(; ゜ 0 ゜)	0.08	0.07	0.07	0.15	0.33	0.39	0.18	**0.95**	0.48	0.32

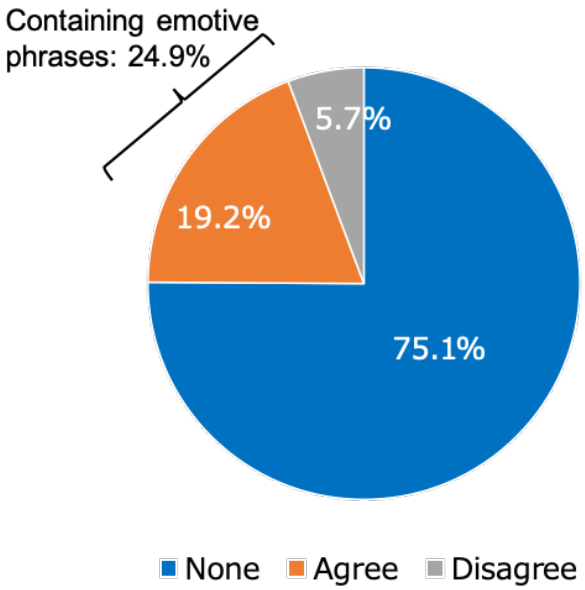


Figure 4.1: Percentage of sentences with emoticons containing emotive phrases [6]. "Agree" indicates that polarity type of an emotive phrase matches emoticon's polarity type within a given sentence.

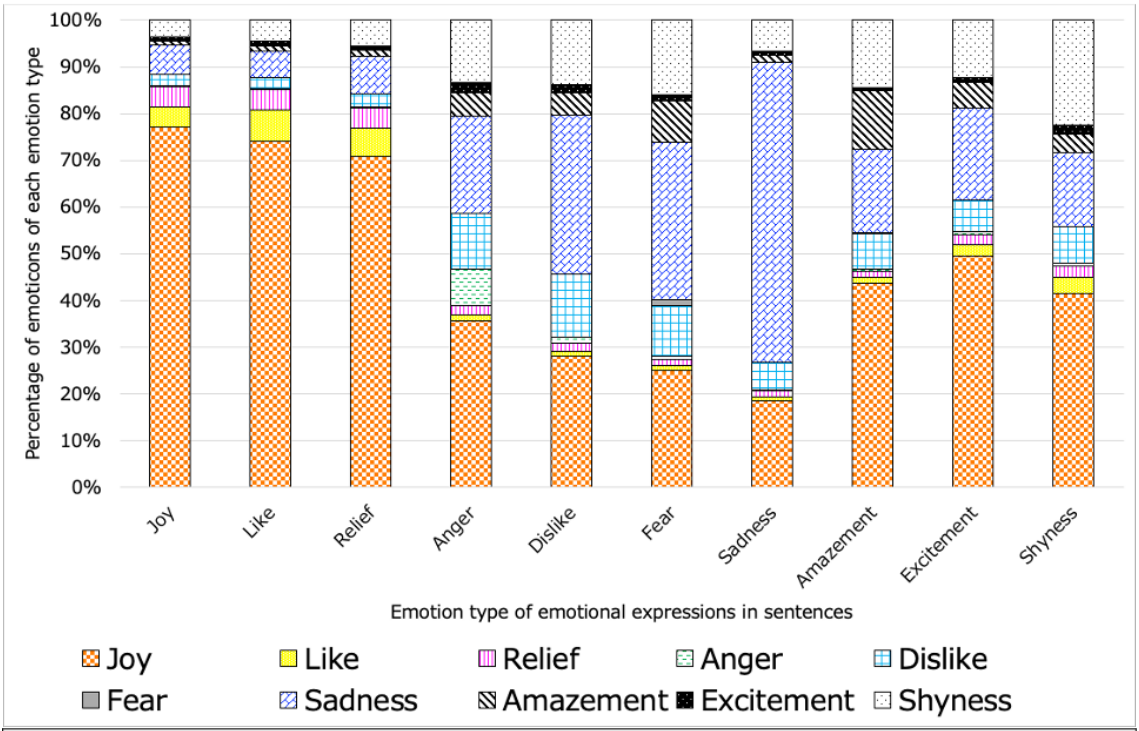
asked to annotate the emotive strength of each ten emotion type in three levels ("weak", "middle", and "strong"). 240 emoticons that expressed higher than "middle" in at least one emotion type were retrieved. Then, nine graduate school students (seven male and two female students) were asked to annotate the strength of each of ten emotion types in five levels to these emoticons in the same way as described in 3.2.2. Finally, the emoticons and their average score of each of ten emotion types were registered in the database. Examples of database entries are shown in Table 4.2. The polarity and the emotion type of the highest value for each emoticon were used to annotate the sentence-emoticon dataset. For example, "(∗^_^∗)" was annotated as POSITIVE for polarity type and JOY for emotion type because JOY had the highest emotive load (see Table 4.2, second row). The emoticon database was used to retrieve sentence-emoticon pair from the aforementioned corpora and collected sentences with 167 out of 299 emoticons. The matched 167 emoticons are listed in Table 4.3.

4.2.2 Results

Figure 4.1 shows the percentage of sentences with emoticons both containing and not containing emotive phrases. 490,345 (75.1%) sentences did not contain any emotive phrases and 162,579 (24.9%) sentences contained emotive phrases. This indicates that the rule-based methods (Chapter **3** and [30, 31]) cannot recommend emoticons for 75.1% of the sentences. For 162,579 (24.9%) sentences containing emotive

Table 4.3: 167 emoticons and their emotion and polarity types

Polarity	Emotion	Quantity	Emoticon				
POSITIVE	JOY	60	(≧ ω≦)	(ﾉ∇≦ *)	ヾ (o´ ∀ ` o)ﾉ	(⌒▽⌒)	☆ (≧ ∀ ≦ *)ﾉ
			Σ d(≧ ∀ ≦ *)	(￣▽￣)	(*≧ m≦ *)	(^。^)	(=^ｪ^=)
			。・゜・ (ﾉ∀ ‘)・゜・。	(^o^)	(*^▽^*)	(#^.^#)	(o^ω^o)
			(*^_^*)	(* ´艸｀)	(○ ´艸｀)	(^ɜ^)-☆	(*ﾉ∀ ‘*)
			(((o(*゜▽゜*)o)))	(^ー゜)	(*^○^*)	(^^)	(≧ 艸≦ *)
			(≧ ɜ≦)	(^O^)	(^ ω ^)	(=^x^=)	＼(^o^)／
			(*^^*)	(ﾍﾄ｀∀ ´) ﾍﾄ	ヾ (● ´ ∀ ` ●)	(^○^)	(^∇ ^)
			(^o^)	(^_^)	(● ` ・ω・)	d(゜∀゜*)	(´艸｀)
			.+(´ ^ ω ^ ‘)+.	(´艸‘*)	(^ O ^)	(● ´ ∀ ` ●)ﾉ	o(≧ ω≦)o
			(` ∀ ´)	(*^o ^*)	(^∀ ^)	(^ω^)	(ﾉ゜∀゜)ﾉ
			(゜∇ ^*)	(^^)	(*´ ∀ ` *)	(^-^)	ヽ (≧ ∀ ≦)ﾉ
			(^人^)	☆*:.。. o(≧ ▽≦)o .。.:*☆		♪———O （≧ ∇≦ ） O———♪	
			(^ ◇ ^)	*・゜ ゜・*:.。..。.:*・’(*゜▽゜*)’・*:.。. .。.:*・゜ ゜・*			
POSITIVE	LIKE	5	(*´v `)	(*б ω б*)	(^_-)	(*v ω v)	(*´ ω ` *)
POSITIVE	RELIEF	9	(u_u)	(*´ɪ‘*)	d(´ ω ` *)	(´ ∀ ` *)	(´ __ `。)
			♪ (´ ▽ `)	(=´ ω ` =)	(o´ ▽ ` o)	(@￣ρ￣@)	
NEGATIVE	ANGER	10	(` Д ´)	(* ` ω ´*)	(− _− #)	(‘皿 ´)	(# ` ∧ ´)
			(゜Д゜#)	o(` ω ´*)o	ヽ (` Д ´)ﾉ	(` ε ´)	(#゜Д゜)
NEGATIVE	DISLIKE	17	(; ￣ｴ￣)	ヽ (´ ρ `)ﾉ	(´ Д‘)	(− __− ;)	(´ Д‘)
			(¬ ɜ ¬)	(~_~;)	(´ Д ` ;)	(￣ー￣)	(´ Д ` υ)
			(--;)	(; ´ Д `)	(; ー_ー)	(ーー ;)	(´A `)
			(￣(工)￣)	(ﾉД‘)			
NEGATIVE	FEAR	1	(((; ゜Д゜)))				
NEGATIVE	SADNESS	27	(´ △ `)	(; _ ;)/~~~	(つД `)	(; д ;)	(/□≦ 、)
			(T__T)	。・゜・ (ﾉД‘)・゜・。	(´・_・‘)	(T_T)	(T . T)
			(T ^T)	(; __ ;)	。゜(゜ ´ Д ` ゜)゜。	ヽ (; ▽ ;) ノ	(>_<)
			(´ ; ω ; `)	(ﾉД‘)	(/ _ ;)	(; ;)	(´・ω・‘)
			。゜゜ (´ □ ` 。)°゜。	。゜(。ﾉω ヽ。)゜。	(σ ´ Д `) σ	。・°°・(>_<)・°°・。	
			il||li(つд-。)il||li	・゜・ (つД `)・゜・	。・゜ ・(ノД‘)・゜ ・。		
OTHER	AMAZEMENT	22	(;゜д゜)	(; ゜ 0゜)	Σ (゜д゜|||)	(゜Д゜)	(・ ・?)
			(((゜Д゜)))	(゜д゜)	Σ (|||▽|||)	(O_O)	! (◎_◎;)
			Σ (゜Д゜)	(゜д゜)	(￣◇￣;)	Σ (・□・ ;)	(゜Д゜)
			(||゜Д゜)	(・・;)	(? _ ?)	Σ (||゜ Д゜)	(; ￣O￣)
			Σ (´ д ` ;)	((((; ゜Д゜))))))			
OTHER	EXCITEMENT	6	(>人<;)	(` ・д・ ´)	(*≧ д≦)	(` ・ω・ ´)	ш (゜Д゜ш)
			Ψ (` ∀ ´) Ψ				
OTHER	SHYNESS	10	(≧ 人≦)	(⌒-⌒;)	(/∀ ＼*)	(〃▽〃)	(^_^;)
			(^◇^;)	(〃ω〃)	(*ﾉɪﾉ)	(*ﾉωﾉ)	(。-_-。)



Number of sentences									
Joy	Like	Relief	Anger	Dislike	Fear	Sadness	Amazement	Excitement	Shyness
68,223 (10.54%)	34,527 (5.29%)	10,017 (1.53%)	1,598 (0.24%)	28,985 (4.44%)	4,576 (0.70%)	7,306 (1.12%)	2,743 (0.42%)	3,296 (0.50%)	1,054 (0.16%)

Figure 4.2: Ratios of emoticons' emotion types included in sentences categorized by emotion [6]. The number of sentences for each emotion type is shown in the table and their percentage in brackets.

phrases, there were 125,362 (19.2%) sentences with selected emoticons expressing the same polarity as the emotive phrases and 37,217 (5.7%) messages containing disagreeing polarities between emoticons and emotive phrases. This means that at least 37,217 (5.7%) inputs cannot be accurately recommended by the rule-based methods which I think that support is needed for these kinds of input sentences.

Figure 4.2 shows the percentage of emotion type of inserted emoticons for each emotion type of sentences. Nearly 20% of sentences containing positive (JOY, LIKE, and RELIEF) expressions did not co-occur with emoticons of the same polarity. For messages containing negative (ANGER, DISLIKE, FEAR, and SADNESS) expressions, 59.3% (ANGER), 51.1% (DISLIKE), 53.6% (FEAR), and 29.7% (SADNESS) of inserted emoticons did not express negative emotion but expressed either positive or other emotion types. For inputs containing AMAZEMENT, EXCITEMENT, and SHYNESS expressions, 72.3% (AMAZEMENT), 81.3% (EXCITEMENT), and 71.7% (SHYNESS) of inserted emoticons expressed either positive or negative emotion. These results suggest that the rule-based methods cannot recommend desired emoticons especially to sentences containing NEGATIVE and OTHER emotive phrases.

Examples of sentence-emoticon pairs with unmatched polarity are shown in Table 4.4. The first example in the table contains a word "笑う" (*warau*, meaning "laugh") which itself has a positive meaning. However, in this context, the sender is being laughed at. Therefore, it can be assumed that the emoticon "(/ _ ;)", which expresses SADNESS, was inserted to express the sender's negative feeling of being ridiculed. The second example in Table 4.4 contains an emotive word "恥ずかしさ" (*hazukashisa*, "shyness" or "embarrassment"). It is then followed by a sentence "消去しました" (*shoukyo shimasita*,

Table 4.4: Examples of sentence-emoticon pairs that their emotion types do not match with each other. Emotive words are emphasized in bold.

Sentence	Emoticon	Sentence		Emoticon	
		polarity	emotion	polarity	emotion
(1) アイコンさっき変えたんだけどみんな**笑う**の (I just changed my avatar but everyone **laughs** about it)	(/ _ ;)	POSITIVE	JOY	NEGATIVE	SADNESS
(2) **恥ずかし**さのあまり消去しました (I deleted it because it was too **embarrassing**)	(; ;)	OTHER	SHYNESS	NEGATIVE	SADNESS
(3) 今日は思いっきり**泣いて**くださいね (You can **cry** as much as you want today)	(*^_^*)	NEGATIVE	SADNESS	POSITIVE	JOY

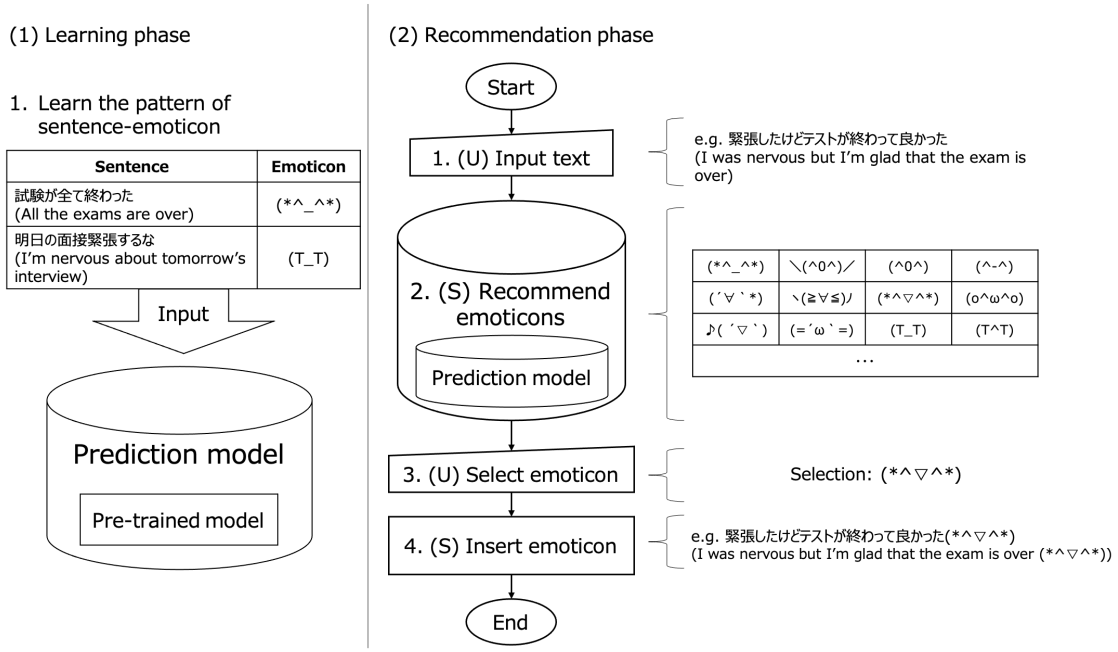


Figure 4.3: System procedure

meaning "I deleted it"), therefore, it can be assumed that the sender had a negative feeling towards the message and added the negative emoticon " (; ;) ". The third example contains an emotive word "泣いて" (*naite*, meaning "cry"), however, this word in the sentence is rather not about the sender but towards someone the sender tries to comfort. Thus, it can be assumed that the sender wants to express a positive feeling to cheer up the receiver.

From the results presented in this section, I concluded that it is not sufficient to recommend emoticons solely from emotive words in a sentence because many inputs do not contain emotive phrases. Moreover, the polarity/emotion types between emotive expressions and emoticons do not always match due to the influence of neighboring non-emotive words. Therefore, I propose a method that can consider the entire information contained in the input message to recommend suitable emoticons.

4.3 Deep-ERS

I propose **Deep-Emoticon Recommendation System** (**Deep-ERS**) to overcome the aforementioned problems: a) cannot recommend emoticons due to the lack of emotive words and phrases in the input sentence and b) an emotion type of the desired emoticon does not match the emotive word/phrase present in the input sentence due to the context of the surrounding words. The distinctive feature of my system is that it incorporates pre-trained models (e.g. Word2vec, ELMo, BERT) that learned the relationship between words in Japanese sentences and utilizes it in the emoticon prediction model to recommend emoticons. The pre-trained models support my system by allowing the use of words appearing in input sentences as a distributed representation which enables to learn patterns between the context of input sentences and chosen emoticons. This is different from the conventional methods that learn patterns between words present in input sentences and selected emoticons using machine learning methods such as support vector machine (SVM) [48]. For example, assume that there are sentences, "I broke my ankle." and "I injured my foot". My system using the pre-trained model can recognize that these sentences are semantically similar to each other, however, the conventional methods only using words present in input sentences as features cannot tell whether they are similar to each other. This means that it can be expected that my system will be able to recommend suitable emoticons to a sentence, "I injured my foot", even if my proposed system only learned the relationship between sentences like, "I broke my ankle" and an emoticon, "(T_T)". In short, the same words do not have to appear in the sentence for which the system is recommending emoticons.

The procedure of **Deep-Emoticon Recommendation System** (**Deep-ERS**) is as follows. The system is divided into two phases: (1) learning phase and (2) recommendation phase. The numbers below correspond to the numbers in Figure 4.3.

Phase (1): learning phase

(1)-1 Learn the pattern of sentence-emoticon pairs. The prediction model uses machine learning techniques (e.g. SVM, LSTM, and fine-tuning). It also incorporates pre-trained models (e.g. Word2vec, ELMo, BERT) learned on large Japanese data.

Phase (2): recommendation phase

(2)-1 A user (U) inputs a sentence.

(2)-2 The system (S) reorders emoticons in the emoticon dictionary according to the result of the prediction model created in the first phase and displays the emoticon list to the user.

(2)-3 The user (U) selects an emoticon from the list.

(2)-4 The system (S) inserts the selected emoticon after the user's input.

Four different methods were tested to discover which one works best: (1) Word2vec [49] vectors with SVM, (2) Word2vec embeddings with attention-based Bidirectional Long Short-Term Memory (BiLSTM) [50, 51], (3) ELMo [52] embeddings with attention-based BiLSTM, and (4) fine-tuning with BERT language model [7]. The details of each pre-trained model are introduced in 4.3.1 and the four aforementioned prediction models are described in 4.3.2. Details of libraries, tools, and pre-trained models

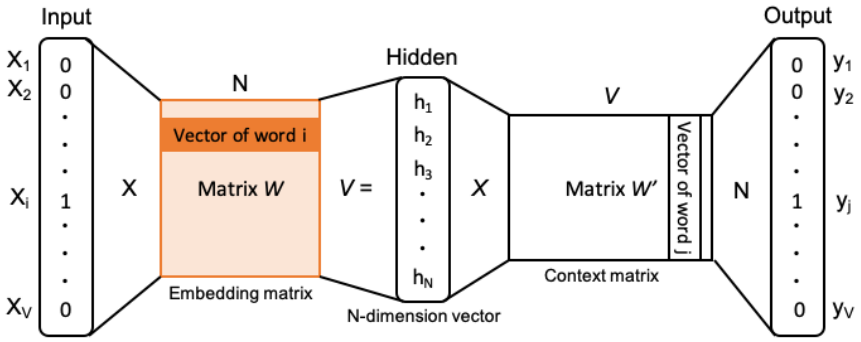


Figure 4.4: Word2vec model [6]

used for the implementation are described in 4.4. The proposed system recommends an ordered set of emoticons as a list. This was based on the hypothesis that there are high chances that various emoticons that express the same emotion type are selected for the same context of sentences.

4.3.1 Pre-trained models

Word2vec

Word2vec [49] is a two-layer neural network that reconstructs the linguistic context of words from natural language inputs (Figure 4.4). The distributed representation of words created by the Word2vec algorithm is used in my prediction models, one combining the Word2vec pre-trained model with SVM classifier and another one using the pre-trained model with attention-based BiLSTM architecture (details are given in Sections 4.3.2 and 4.3.2). It learns linguistic context using either of the following two methods: (1) continuous bag-of-words (CBOW) which learns the surrounding words to predict the word in the middle and (2) Skip-gram which predicts the surrounding words from the input word. By learning natural language text with the Word2vec algorithm, a distributed representation of words is retrieved as vectors ("Embedding matrix" in Figure 4.4). The vectors of words that are used in a similar context share similar directions. Figure 4.4 presents the Skip-gram model to explain the "Embedding matrix" which is used in the prediction models. Both the input vector $(x_1, x_2, ..., x_V)$ and the output vector $(y_1, y_2, ..., y_V)$ are one-hot encoded word representations. The hidden layer is the word embedding of size N which can be set manually. The "Embedding matrix" in the figure is the multiplication of the input vector X and the matrix W of size $N \times V$. V is the size of the vocabulary. Matrix W is determined by learning the semantic relationships between words based on the surrounding words in a sentence. "Vector of word" in the figure is the distributed representation for the i-th word in the vocabulary. "Context matrix" is the multiplication of the hidden vector and the matrix W' of size $V \times N$. Matrix W' is optimized to predict the surrounding words for the target input word. By using the "Embedding matrix", a static embedding vector for each word can be retrieved. The reason to use the Word2vec pre-trained model originates from the hypothesis that semantically similar texts tend to co-occur with similar emoticons and the Word2vec can be helpful to grasp the semantic similarity of sentences and predict more acceptable emoticons.

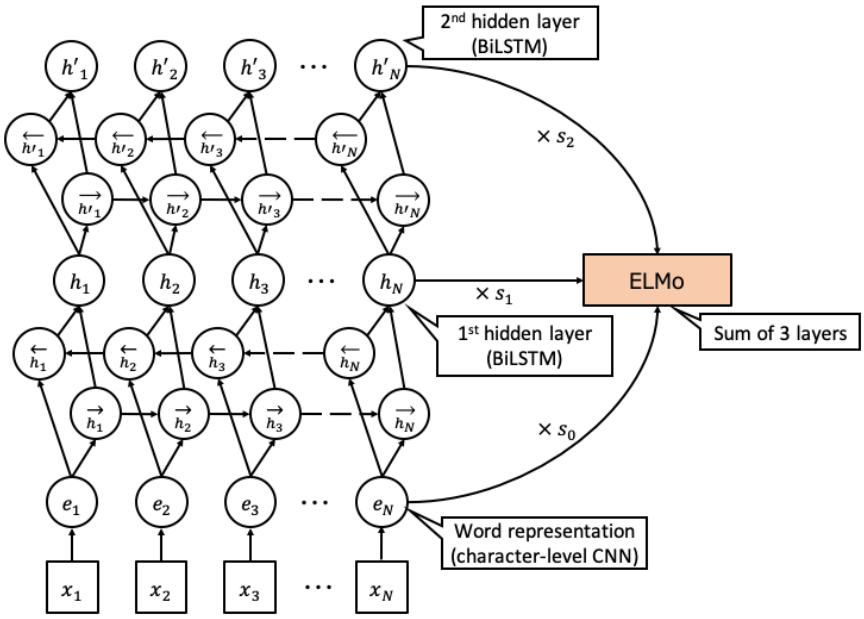


Figure 4.5: ELMo model [6]

ELMo

The ELMo pre-trained model can output word representations that take the entire sentence or paragraph into account that is different from the Word2vec model which can only map a word to a fixed vector. Word embedding vectors retrieved from ELMo are a weighted sum of three layers; a) word representation vectors created using a character-level convolutional neural network [53] from a text string input, b) a hidden layer that is a concatenation of forward and backward information (BiLSTM) created from the word representation vectors, and c) another BiLSTM layer created from the previous hidden vector (Figure 4.5). s_0, s_1, and s_2 in the figure are weights for each layer. ELMo pre-trained model was employed expecting to capture the context of an entire input sentence to recommend emoticons. The pre-trained model was applied in my prediction model with attention-based BiLSTM (details are given in 4.3.2).

BERT

BERT [7] is a multi-layer bidirectional transformer encoder that is trained on masked word prediction and next sentence prediction. Transformer [54] learns positions of words in an input sentence and uses a multi-head attention mechanism to learn contextual relations between words in a text. As described in [7], BERT randomly replaces words with a special token ("mask") in a sentence and learns the context of input sentences by predicting the masked word based on the unmasked surrounding words, both left and right, using a transformer. This allows learning deeper than ELMo which only uses the concatenation of independently trained left-to-right and right-to-left LSTM. The fine-tuning approach using BERT was employed in **Deep-ERS** because this model has achieved remarkable results in many natural language understanding tasks. Also, I expect that this model could recommend emoticons regarding the contextual relations between words in an input text. The fine-tuning approach using BERT is described in 4.3.2.

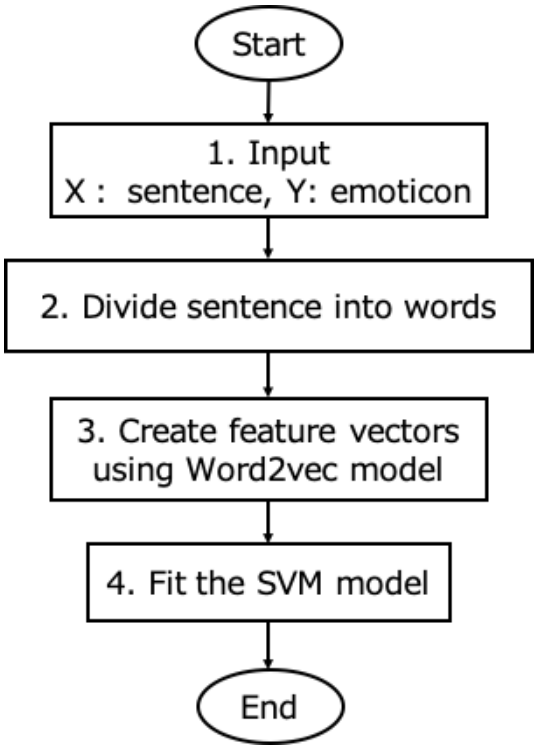


Figure 4.6: Outline of **Word2vec+SVM** model [6]

4.3.2 Prediction models

Word2vec+SVM

Prediction model using vectors retrieved from Word2vec and SVM classifier (hereinafter referred to as **Word2vec+SVM**) is shown in Figure 4.6. The following numbers correspond to the numbers in that figure.

1. Input sentence-emoticon pair as training data.
2. Divide sentences into words (tokenize).
3. Create feature vectors. The feature vector for the input sentence is the average of each word vector retrieved from the Word2vec model. It can be represented with Equation 4.1. In this equation, assume that there are N words (w_1, w_2, ..., w_N) in a sentence s. For each word, a word embedding vector ($v_{w_i}, i = 1, 2, ..., N$) is retrieved from the Word2vec model. Then, the average vector is calculated by dividing the sum of the word embedding vectors by the number of words (i.e. N) in the sentence. Only words that exist in the Word2vec model was used.

$$V_s = \frac{\Sigma_{i=1}^{N} v_{w_i}}{N} \tag{4.1}$$

4. Fit the training data to the SVM model.

Word2vec/ELMo+BiLSTM+Attention

BiLSTM [50] is one of the recurrent neural network architectures which is composed of two LSTMs: one taking the input in a forward direction, and the other in a backward direction. This method is

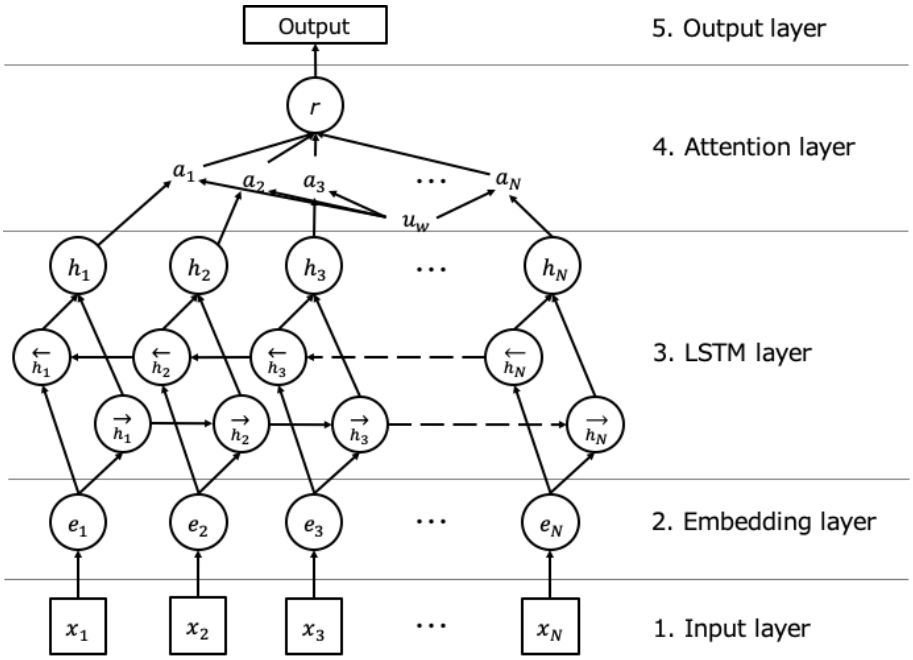


Figure 4.7: Word2vec/ELMo+BiLSTM+Attention model [6]

capable of learning the past and future information in the sequence. Attention mechanism [51] was also applied to learn important words in the input sentence to recommend suitable emoticons. The reason attention-based BiLSTM with Word2vec or ELMo models was employed as proposed methods was that this combination has yielded good results in NLP tasks (e.g. emoji prediction and sentiment analysis) [52, 55]. I also considered that the prediction result would be better if a method not only learns the sequence of words in sentences but also focus on important words to recommend emoticons. The description of the model using either pre-trained Word2vec model or ELMo model for word embeddings with attention-based BiLSTM (hereinafter referred to as **Word2vec+BiLSTM+Attention** and **ELMo+BiLSTM+Attention**) is shown in Figure 4.7). The numbers below correspond to the numbers in the figure.

1. Input layer: input sentences to the model (sentences are tokenized).

2. Embedding layer: convert the input vector to a matrix containing each word embedding vector retrieved from the Word2vec or the ELMo model. For the Word2vec pre-trained model, a static word embedding vector for each word in the input sentence was obtained. For the ELMo model, the entire input sentence is used as an input to the pre-trained model, and a weighted sum of the three layers (see 4.3.1) for each word was retrieved.

3. LSTM layer: input the embedding layer into BiLSTM layer. BiLSTM layer consists of forward hidden state $\overrightarrow{h_i} \in \mathbb{R}^{d_h}$ and backward hidden state $\overleftarrow{h_i} \in \mathbb{R}^{d_h}$, where d_h denotes the number of the hidden units. Both layers (i.e. $h_i = [\overrightarrow{h_i}; \overleftarrow{h_i}] \in \mathbb{R}^{2d_h}$) are concatenated.

4. Attention layer: input the LSTM layer into the Attention layer. The attention layer implemented in this architecture is described in [51] which is also used in other NLP classification tasks and has shown its effectiveness [51, 56, 57]. The attention layer calculates the representation r using three equations shown below (Equations 4.2-4.4). First, the hidden state h_i is fed through a

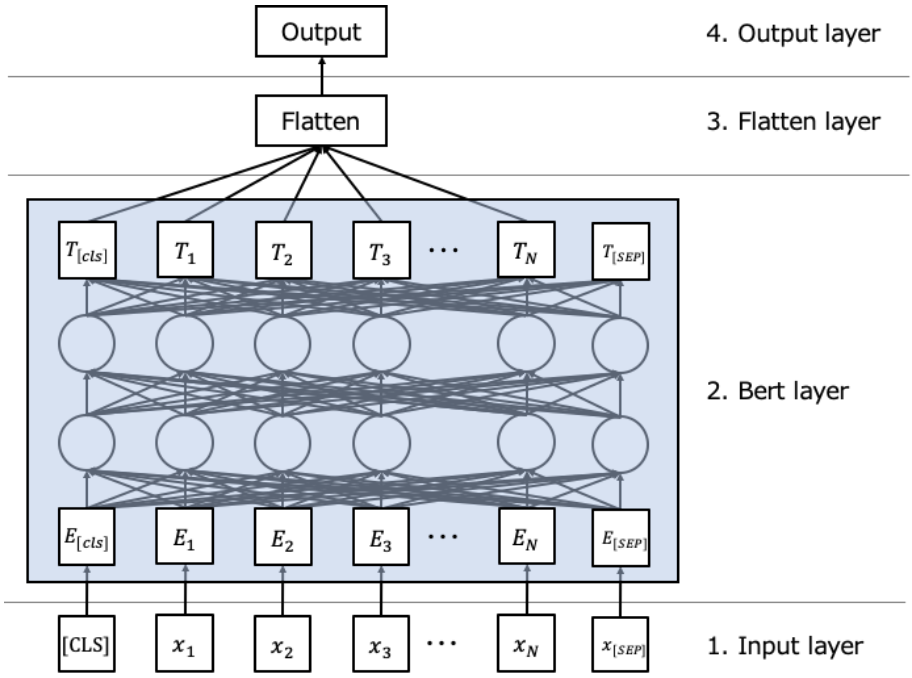


Figure 4.8: BERT+fine-tuning model [6]

one-layer Multi-Layer Perceptron to obtain hidden representation u_i (Equation 4.2). W_w and b_w are weight and bias, respectively, and are optimized during the training period. Then, the weight a_i for h_i is computed as the similarity between u_i and a word-level context vector u_w (Equation 4.3). The calculation for weight uses the softmax function to convert the similarity (u_i and u_w) into a normalized probability distribution. The context vector u_w is considered as a high-level representation to judge the importance of words in a sentence [51]. The context vector can be considered as a "weight" used with the hidden representation u_i. It is randomly initialized to allow different weights to learn independently of each unit and also jointly learned during the training process. Finally, the representation r is computed as a weighted sum of the hidden state vectors with weights (Equation 4.4).

$$u_i = tanh(W_w h_i + b_w) \tag{4.2}$$

$$a_i = softmax(u_i^T u_w) = \frac{exp(u_i^T u_w)}{\Sigma_{i=1}^{N} exp(u_i^T u_w)} \tag{4.3}$$

$$r = \Sigma_{i=1}^{N} a_i h_i \tag{4.4}$$

5. Output layer: the output is decoded to a probability of each emoticon using the softmax function.

BERT+fine-tuning

The description of the fine-tuning approach using BERT (hereinafter referred to as **BERT+fine-tuning**) is given below. The numbers correspond to the numbers in Figure 4.8.

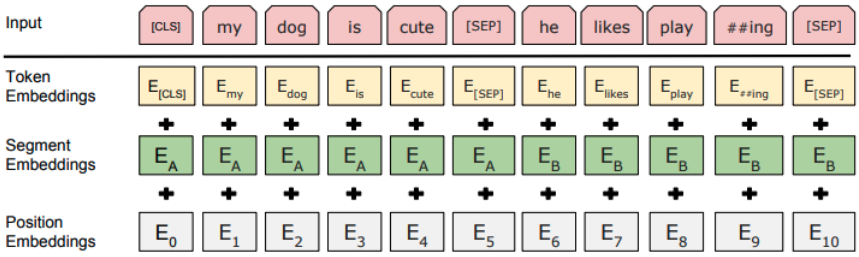


Figure 4.9: BERT input representation [7]

1. Input layer: input a sentence into the model. The sentence is tokenized and "[CLS]" is added before the first token and "[SEP]" is added after the last token as explained in [7] (See Figure 4.9).

2. BERT layer: convert the input layer to the BERT layer. The input embeddings ($E_{[cls]}$, E_1, E_2, ..., $E_{[sep]}$) are the sum of the token embeddings, the segmentation embeddings, and the position embeddings (Figure 4.9). The matrix size of the BERT layer output is the number of tokens in the sentence, each having 768 dimensions. For example, if the number of tokens in an input sentence is 20, then the matrix size of the output is 20×768. The number 768 is a default number of dimensions used in [7].

3. Flattening layer: convert the BERT layer into a flatten layer to convert the matrix into a single array. For example, if the matrix of the BERT layer is 20×768, then it is transformed into a single array containing 15,360 ($= 20 \times 768$) dimensions.

4. Output layer: the output is decoded into a probability of each emoticon using the softmax function.

4.4 Evaluation experiment

Two evaluation experiments, (A) recommendation accuracy and (B) Mean Average Precision (MAP) [58], were performed to discover which prediction model (see 4.3.2) recommends desired emoticons best. The details are given in 4.4.1. The dataset used in the experiment is explained in 4.4.2. The proposed methods were compared with the conventional prediction methods that learned from the same sets of surface patterns of sentences and emoticons. The details are described in 4.4.3. Finally, the detailed implementation of the proposed prediction algorithms using pre-trained models is explained in 4.4.4.

4.4.1 Evaluation methods

(A) Recommendation accuracy

The recommendation accuracy is computed by counting the number of correct emoticons that were recommended within the top 25. By "correct emoticons", I mean all facemarks used in sentences of the test data (details of the test data are described in 4.4.2). The number of top emoticons was set to 25 because the number of emoticons shown on a single smartphone screen (before scrolling down) is about 25 emoticons (the number changes depending on the length of emoticons). The evaluation process can

determine whether a given prediction model is effective when more desired emoticons are recommended among the first 25 emoticons which users see on their screens at the first glance without scrolling down.

Recommendation accuracy was calculated by using Equation 4.5. $A@N$ denotes the accuracy of correct emoticons recommended within the top N. For example, $A@5$ means the percentage of correct emoticons recommended was within the top 5 output candidates. S is the number of samples and r_i is the ranking of the correct emoticon for i-th sample.

$$A@N = \frac{\Sigma_{i=1}^{S} f(r_i)}{S} \times 100(\%), f(r_i) = \begin{cases} 1, & \text{if } r_i \leqq N. \\ 0, & \text{otherwise.} \end{cases} \tag{4.5}$$

(B) Mean Average Precision (MAP)

MAP [58] is a method common in the recommendation technology field which evaluates how many of the relevant items are recommended as the top candidates from the entire pool of recommended items. MAP of each prediction model was evaluated to compare which model is likely to recommend emoticons of the same polarity or the emotion type of the correct emoticon together at the top of the list. As users tend to choose emoticons of the same polarity/emotion type for similar contexts of texts, it can be considered that the prediction model is effective if emoticons expressing similar sentiment are recommended with the correct emoticon at top of the list. "MAP_polarity", indicating how emoticons of the same polarity are recommended, and in the case of "MAP_emotion", performing the same evaluation but focusing on the emotion type were evaluated. However, the results for OTHER, namely AMAZEMENT, EXCITEMENT, and SHYNESS type emoticons when calculating MAP_polarity were excluded because AMAZEMENT, EXCITEMENT, and SHYNESS are not as similar to each other (also they can be either positive or negative).

MAP was calculated by using Equations 4.6 and 4.7. In Equation 4.6, S denotes the number of samples. In Equation 4.7, M is the number of desired emoticons and $Precision@k \cdot y_k$ indicates the precision of relevant items within top k. For example, suppose that relevant items are in rank 1 and rank 3. AP is then $\frac{1}{2} \times (\frac{1}{1} + \frac{2}{3}) = 0.83$ (point). The MAP value close to 1 point means the prediction model predicts relevant emoticons at the top of the list and a value close to 0 point means that relevant emoticons are recommended at the bottom of the list.

$$MAP = \frac{1}{S} \Sigma_{s \in S} AP(s) \tag{4.6}$$

$$AP(s) = \Sigma_{k=1}^{M} \frac{Precision@k \cdot y_k}{\Sigma_{i=1}^{k} y_i}, y_k = \begin{cases} 1, & \text{emoticon ranked within top } k. \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

4.4.2 Data description

The data described in 4.2 was imbalanced in terms of the number of sentences containing each emoticon. Therefore, a maximum of 2,000 data samples for each emoticon for training data and 400 data samples per emoticon for test data were collected from the dataset used in the preliminary investigation (see 4.2). Moreover, the data samples were retrieved randomly following the procedure shown in Figure 4.10 to avoid the imbalance in terms of the emotion type of emotive phrases contained in text messages including sentences without emotive phrases. This procedure is more complex than just randomly retrieve

Table 4.5: Number of sentences for each emotion type (training data)

Emotion type	JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SADNESS	AMAZEMENT	EXCITEMENT	SHYNESS	NONE
Number of sentences	11,600	7,935	4,464	822	9,046	2,196	2,658	1,470	1,730	566	81,872
Total	124,359										

Table 4.6: Number of sentences for each emotion type (test data)

Emotion type	JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SADNESS	AMAZEMENT	EXCITEMENT	SHYNESS	NONE
Number of sentences	3,867	2,696	1,450	639	3,729	1,319	1,424	901	1,043	480	20,074
Total	37,622										

data samples but was employed because the whole dataset contained 75% of sentences without emotive expressions and may result in collecting more NONE type sentences and few samples of sentences expressing emotive content.

As shown in Figure 4.10, there are 42,203 samples for emoticon "(^_^;)" in the dataset. If the samples are collected randomly, it may lead to gathering more samples categorized as NONE and fewer samples from the other emotion types because 80% of the samples are categorized as NONE. Therefore, to create training and test data, first, the number of samples is checked whether it exceeds 2,400, the biggest set of samples for a single emoticon in the training and test data. If the number of the samples is more than 2,400, then the number of samples for retrieval is computed by dividing 2,000 by the number of existing emotion types for the emoticon including NONE. For example, as shown in Figure 4.10, samples exist for all emotion types, therefore, 2,000 is divided by 11, and 182 (= 2000/11) is determined as the number of samples to extract. Then, the samples are retrieved from the emotion type that contained the smallest number of samples (Figure 4.10, SHYNESS, 147 samples). If the number of samples is smaller than the number of extraction, only 50% of the samples are collected so the rest of the samples can be used in the test data. In Figure 4.10, 73 sentences are collected for SHYNESS which is nearly 50% of the samples. Next, the number of retrievals is recalculated by dividing the remaining total number of samples with the rest of the emotion types. For example, the number of samples to retrieve changed to 193, dividing 1,927 samples ($= 2,000 - 73$) by the rest of the ten emotion types. Then, the samples are collected according to the emotion type containing the second smallest samples (ANGER, 179, see Figure 4.10). This procedure is repeated until samples are retrieved from all categories. The procedure to create the test data was the same but the maximum number of retrieval for each emoticon was set to 400. If the number of samples for the emoticon is less than 2,400, 50% of samples for each emotion type was used in the training data and the rest in the test data. The size of the training data was 124,359 samples containing 167 different emoticons. The details are shown in Table 4.5. The size of the test data was 37,622 samples containing the same number of different emoticons in the training data (Table 4.6).

4.4.3 Baseline methods

Comparison of **Deep-ERS** and the four conventional methods was conducted to investigate the effectiveness of the proposed methods. The four conventional methods are described below. I could not use my previously proposed method **EBReS-η** for comparison because there is no user information in the

Total	Number of sentences containing emotional phrases of emotion type for emoticon "(^_^;)"					
	Joy	Like	Relief	Anger	Dislike	Fear
42,203	1,769	848	423	179	3,434	615
	Sadness	Amazement	Excitement	Shyness	None	
	382	332	316	147	33,758	

1.Extract samples for training data

"(^_^;)": 42,203 > 2,400

$Number\ of\ samples\ to\ retrieve = \frac{2{,}000}{11} \approx 182,$

Shyness: 147 < 182
Retrieve 50% of the samples from 147 samples.
Shyness: 73 ($\approx 50\%$)

$Number\ of\ samples\ to\ retrieve\ = \frac{2{,}000-73}{10} \approx 193,$

Anger: 179 < 193
Retrieve 50% of the samples from 179 samples. Anger: 89 ($\approx 50\%$)

$Number\ of\ samples\ to\ retrieve\ = \frac{2{,}000-(73+89)}{9} \approx 204,$

Excitement: 316 > 204
Retrieve 204 samples from sentences containing Excitement phrases.

Repeat the process for the remaining emotion types (i.e. Amazement, Sadness, Relief, Fear, Like, Joy, Dislike, and None).

2. Extract samples for test data

Exclude samples registered in the training data.

"(^_^;)": 40,203 > 400

$Number\ of\ samples\ to\ retrieve = \frac{400}{11} \approx 36,$

Shyness: 74 > 36
Retrieve 36 samples from 74 samples.

Repeat the process for the remaining emotion types.

Total	Number of sentences in training data for emoticon "(^_^;)"					
	Joy	Like	Relief	Anger	Dislike	Fear
2,000	204	204	204	89	205	204
	Sadness	Amazement	Excitement	Shyness	None	
	204	204	204	73	205	

Total	Number of sentences in test data for emoticon "(^_^;)"					
	Joy	Like	Relief	Anger	Dislike	Fear
400	37	37	36	36	37	36
	Sadness	Amazement	Excitement	Shyness	None	
	36	36	36	36	37	

Figure 4.10: Procedure for creating training/test data samples [6]

dataset[4]. Instead, I employed a method similar to Emura *et al.*'s method [31] which analyzes the emotion type of the input sentence and recommends frequently used emoticons.

Frequency method

The **Frequency** method recommends emoticons based on the number of times selected in the past by all users. This method was employed because I believe that this method is a similar algorithm to Emura *et al.*'s method [31]. The frequency of users' selection for each emoticon was computed using the training data. More concretely, the frequency for each emoticon is counted while taking the emotion type of the emotive phrases in the input sentence into account. For example, if the emotive expression of the sentence belongs to the JOY type, only sentences containing JOY phrases were used to count the number of selections for each emoticon and recommended emoticons in descending order. However, if no emotive expression exists in the sentence, the system recommended frequent emoticons from the NONE type sentences in the training data.

BoW+SVM method

The **BoW+SVM** method recommends emoticons by learning the patterns of sentences and emoticons using Bag of Words (BoW) as features with SVM classifier [36]. This method was employed because it is widely used in various NLP tasks as a baseline method. To create a feature vector of a sentence, it first examines the number of the vocabulary of a whole data and creates a vector, each element representing each word. Then, it counts the number of each word appearing in a sentence and maps the values to the corresponding dimension in a vector. Let us see two sentences shown in Figure 4.11: "This movie is very scary and very long." and "This movie is long but exciting". The size of the vocabulary is 9 ("This", "movie", "is", "very", "scary", "and", "long", "but", "exciting") so the vector of 9 elements is created. The feature vector of the first sentence is [1, 1, 1, 2, 1, 1, 1, 0, 0], expressing the number of occurrences of each word in the sentence. Concretely, words "This", "movie", "is", "scary", "and", "long" are expressed as 1 in the feature vector because they only appear once in the sentence. The word "very" is expressed as 2 in the feature vector because "very" appears twice in the sentence. Words "but" and "exciting" don't exist in the sentence so they are expressed as 0 in the feature vector. The feature vector is then fed to the SVM classifier.

The Scikit-learn [59] library was used for the implementation. C-Support Vector Classification (C-SVC) with Radial Basis Function (RBF) kernel was employed. The hyperparameters, gamma (γ), and C were determined by performing grid search ranging from 0.01 to 10 and used the parameters that yielded the best A@25 on the test data. The gamma parameter defines how strong each training sample influences the prediction. The C parameter is a trade-off of correct classification of training examples against maximization of the decision function's margin[5]. The same way of using SVM was applied for other baseline methods **TF-IDF+SVM** and **WordCharGram+SVM** and the proposed method **Word2vec+SVM** which is explained later. The hyperparameters used for **Bow+SVM** were 0.1 for γ and 10 for C (see Table 4.7). The system sorted the emoticons in an order based on the value of the

[4]**EBReS-η** recommends emoticons based on the user's emoticon selection history and the emotion type of an emotive phrase recognized in input sentences. As the dataset does not include user information, the system cannot recommend emoticons using the user's emoticon selection history.

[5]For details, see `https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html`.

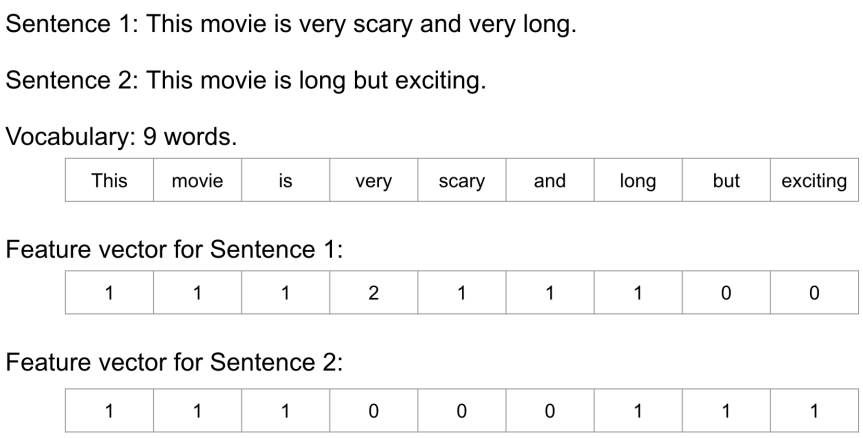


Figure 4.11: Examples of creating Bag of Words feature vectors from sentences

decision function which gives the probability of the prediction for each emoticon. Since the dataset was in Japanese in which words are not divided by spaces, a morphological analysis tool MeCab [60] was used for tokenization. MeCab was also used in **TF-IDF+SVM** and **WordCharGram+SVM** methods.

TF-IDF+SVM method

The **TF-IDF+SVM** method is basically the same as **BoW+SVM** but uses TF-IDF [61] score of words as features. This method was employed because it is also popularly used in various NLP tasks as a baseline method. The procedure for creating a feature vector of a sentence is mostly the same as creating it from BoW but the value for each word is calculated using the TF-IDF method. The TF-IDF method was proposed by [61] and it is intended to reflect how important a word is to a document in a document collection. The equations for TF-IDF are shown in Equations 4.8, 4.9, and 4.10. In Equation 4.8, w indicates a word and s stands for a sentence. In Equation 4.9, $f_{w,s}$ denotes the number of words w in the sentence s. In Equation 4.10, N denotes the number of sentences in a collection and n_w means the number of sentences that contain the word w. Based on these equations, words that do not appear in many sentences and only exist in certain sentences have more importance in prediction than the words that appear in many sentences. The feature vectors were also fed to the SVM classifier using hyperparameters, 0.1 for γ and 10 for C (see Table 4.7).

$$TF - IDF_{w,s} = TF_{w,s} \times IDF_w \tag{4.8}$$

$$TF_{w,s} = f_{w,s} \tag{4.9}$$

$$IDF_w = log(\frac{N}{1 + n_w}) + 1 \tag{4.10}$$

WordCharGram+SVM method

The **WordCharGram+SVM** method uses word and character n-gram as features with the SVM classifier [36]. This method was employed because it worked best in SemEval emoji prediction task [3]. The creation of the feature vector is shown in Figure 4.12. In this figure, there is a sentence “This movie is

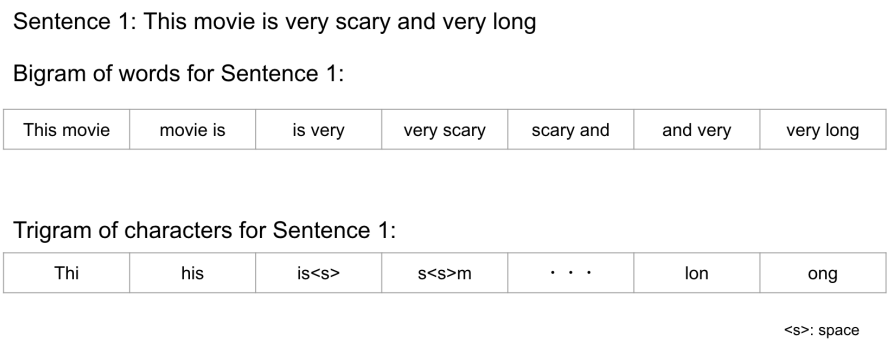


Figure 4.12: An example of creating a feature vector for a sentence with **WordCharGram**

Table 4.7: The best parameters for each baseline method

Method	γ	C
BoW+SVM	0.1	10
TF-IDF+SVM	0.1	10
WordCharGram+SVM	0.1	10

very scary and very long." Assume that the n-gram for words is set to 2 (bigram) and the n-gram for characters is set to 3 (trigram). Then, the bigram list of words become, "This movie", "movie is", "is very", ..., "very long". The trigram list of characters will be, "Thi", "his", "is<s>"[6], ..., "lon", "ong". These word bigrams and character trigrams all together form one vector in which each of them is mapped as an element. Word bigrams and character bigrams were used as features, respectively, as they achieved the best A@25 among 2-5 grams. TF-IDF method was applied for scoring each element. 0.1 for γ and 10 for C were set as hyperparameters for SVM (see Table 4.7).

4.4.4 Implementation for each proposed prediction model

Word2vec+SVM

Gensim [62] library was employed to load the pre-trained Word2vec model and Scikit-learn library to implement SVM. Two pre-trained Word2vec models were used to create features from sentences: **Word2vec (Wiki)**, a 300-dimensional model learned from Japanese Wikipedia articles [63] and **Word2vec (SNS)**, a 200-dimensional model learned from social media corpus [64]. These two pre-trained models were used separately to examine whether the difference in the learning domain (Wikipedia or social media) influences the recommendation results. Since emoticons are more likely to be used in casual communication,

[6]"<s>" indicates "space"

Table 4.8: The best parameters for each prediction model (models using SVM) and number of dimensions of pre-trained models.

Method	γ	C	Dimension
Word2vec (Wiki)+SVM	1.0	10	300
Word2vec (SNS)+SVM	0.01	10	200

Table 4.9: The best parameters for each prediction model (models using deep learning techniques)

Method	Dimension	LSTM unit	Batch size	Epoch	Learning rate
Word2vec (Wiki)+BiLSTM+Attention	300	200	16	13	$1e$-4
Word2vec (SNS)+BiLSTM+Attention	200	200	16	8	$1e$-4
ELMo+BiLSTM+Attention	1024	200	16	4	$1e$-5
BERT (Wiki)+fine-tuning	768	-	16	4	$1e$-5
BERT (SNS)+fine-tuning	768	-	16	18	$1e$-6

using a pre-trained model trained on social media corpus may achieve better recommendation results than learning from non-casual information like Wikipedia articles. MeCab was used for tokenization to match the way of tokenization performed in the pre-trained Word2vec models. Words that do not exist in the pre-trained models were excluded. 87.8% and 78.7% of all words in both training and test data existed in Word2vec (SNS) and Word2vec (Wiki), respectively. The hyperparameters, γ and C that yielded the best A@25 with test data were used. For the model using **Word2vec (Wiki)**, γ and C were 1.0 and 10, respectively (see Table 4.8, **Word2vec (Wiki)+SVM**). For the model using **Word2vec (SNS)**, gamma and C were 0.01 and 10, respectively (see Table 4.8, **Word2vec (SNS)+SVM**). The system sorted the emoticons in order based on the value of the decision function which gives the probability of the prediction for each emoticon.

Word2vec+BiLSTM+Attention

Tensorflow Keras library[7] was employed to create a deep learning-based model. Tensorflow Keras was also used for implementing **ELMo+BiLSTM+Attention** and **BERT+fine-tuning**. MeCab was also used here to tokenize texts for the same reason explained in 4.4.4. As deep learning methods using sequential data require fixed length, the last 20 tokens for each sentence were used. The number was set to 20 tokens because 94.2% of the sentences contained less than 20 tokens. The "last" 20 tokens were used for the prediction because the latter words have more influence when choosing emoticons [31]. For the embedding layer, the same Word2vec models described in 4.4.4 were used. For the BiLSTM layer, 200 was set for units (100 units for each forward and backward processing) and the batch size was set to 16. The learning rate was determined by investigating which learning rate among 0.01 to $1e$-6 yields the best A@25 for the test data. The best learning rate was $1e$-4 (see Table 4.9, **Word2vec (Wiki)+BiLSTM+Attention** and **Word2vec (SNS)+BiLSTM+Attention**). The learning ran until 20 epochs and the number of epochs that yielded the best A@25 for the test data was 13 for **Word2vec (Wiki)+BiLSTM+Attention** and 8 for **Word2vec (SNS)+BiLSTM+Attention**. Adam [65] was applied as an optimizer which is commonly used and known as effective in natural language-based tasks. Categorical cross-entropy was used as the loss function which is calculated by Equation 4.11. In this equation, y is an indicator giving either 0 or 1 indicating whether a class label is correctly classified, and P is the predicted probability of the correctly classified label. $i \in [1, S]$ is the index of training samples and $j \in [1, N]$ is the index of the class label. Adam and categorical cross-entropy are also used in **ELMo+BiLSTM+Attention** and **BERT+fine-tuning**.

[7] Tensorflow-keras: `https://www.tensorflow.org/guide/keras`, version 2.1.0-rc2.

$$Loss = -\Sigma_{i=1}^{S}\Sigma_{j=1}^{N}(y_{ij}log(P_{ij})) \tag{4.11}$$

ELMo+BiLSTM+Attention

ELMoForManyLangs library[8] was employed to retrieve word vectors to be used in the embedding layer from the pre-trained ELMo model. The pre-trained ELMo model[9] which was trained on 400,000 Japanese business news articles was used as it was the only Japanese ELMo model freely available. MeCab was also employed for tokenization to match the text chunks in the pre-trained ELMo model. For the input layer, the last 20 tokens for each input sentence was used, which is as same as in 4.4.4. 200 units for the BiLSTM layer and the batch size of 16 were set for learning. The learning rate and the number of epochs for the model used for the comparison were $1e$-5 and 4, respectively (see Table 4.9, **ELMo+BiLSTM+Attention**).

BERT+fine-tuning

Keras-for-tf2 library[10] was employed to load and fine-tune the pre-trained BERT model. Two pre-trained BERT models were used: **BERT (Wiki)**, a model trained on Japanese Wikipedia articles [66] and **BERT (SNS)**, a model trained on social media corpus [67]. SentencePiece [68] was utilized for tokenization. For the input layer, the last 20 tokens were used. The learning rate and the number of epochs for the model used for the comparison were $1e$-5 and 4, respectively, for the **BERT+fine-tuning** model using pre-trained **BERT (Wiki)** model (see Table 4.9, **BERT (Wiki)+fine-tuning**). The learning rate and the number of epochs for the model used for the comparison were $1e$-6 and 18 epochs, respectively, for BERT+fine-tuning model using **BERT (SNS)** model (see Table 4.9, **BERT (SNS)+fine-tuning**). The batch size was set to 16.

4.5 Results and discussion

4.5.1 (A) Recommendation accuracy

Table 4.10 shows the recommendation accuracy of correct emoticons recommended within the top 25. The results were also divided according to the emotion type of emotive expression in a sentence to see whether there is a difference in the accuracy between sentences that contain an emotive expression of a certain emotion category including NONE type sentences. Wilcoxon signed-rank test [69] was applied to investigate a statistical difference between rankings of the correct emoticon for each sample given by the prediction models (see Table 4.11). Wilcoxon signed-rank test is a non-parametric statistical test to compare two related samples to assess whether the population mean ranks differ. A P value lower than 0.05 commonly indicates that the two samples are significantly different from each other. The P value of all pairs except for one pair, **BoW+SVM** & **TF-IDF+SVM** were lower than 0.01 indicating that there are significant differences in the recommendation results outputted by the prediction models. The P value for the pair **BoW+SVM** & **TF-IDF+SVM** was 0.06.

[8]ELMoForManyLangs: `https://github.com/HIT-SCIR/ELMoForManyLangs`, version 0.0.3.
[9]ELMo model: `https://stockmark.co.jp/news/2019/01/25/783/`, created by a Japanese company, Stockmark Inc.
[10]Keras-for-tf2: `https://github.com/kpe/bert-for-tf2/`, version 0.13.5.

Table 4.10: Recommendation accuracy (A@25) (%). Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

Method	Average	Emotion type of emotive expression in inputs (with number of samples)										
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SADNESS	AMAZEMENT	EXCITEMENT	SHYNESS	NONE
	A@25	1,930	1,494	942	420	2,047	812	947	588	720	324	9,708
(1)	29.5	24.4	32.5	52.7	**68.2**	32.2	56.4	60.5	61.4	59.7	**70.6**	18.8
(2)	42.4	40.1	45.5	56.8	56.3	45.5	55.1	54.6	55.3	57.1	60.4	36.8
(3)	42.5	41.1	46.8	57.0	55.7	45.2	55.5	54.9	56.9	59.1	60.0	36.5
(4)	43.6	43.6	49.6	59.9	55.9	46.6	55.8	54.0	58.4	59.4	60.8	37.2
(5)	44.2	43.4	49.3	57.8	56.7	45.8	53.4	55.7	58.3	59.3	57.5	38.8
(6)	45.6	44.5	50.6	60.0	59.3	48.6	56.3	57.5	58.8	60.0	58.5	39.9
(7)	47.6	43.8	51.6	62.0	59.2	50.0	60.7	63.1	62.4	63.6	65.6	42.1
(8)	48.4	44.6	49.1	60.8	62.4	50.0	60.3	64.5	63.3	65.2	66.5	43.5
(9)	47.4	44.5	50.3	59.0	59.6	48.8	57.7	60.7	62.9	61.7	58.3	42.8
(10)	50.4	47.3	51.5	62.5	62.0	50.2	**62.5**	64.2	63.6	64.7	65.6	46.2
(11)	**53.0**	**49.9**	**55.4**	**65.0**	65.7	**54.9**	61.6	**66.5**	**65.3**	**69.0**	67.5	**48.4**

As can be observed from Table 4.10, **BERT (SNS)+fine-tuning** achieved the highest average accuracy of 53.0% and **Frequency** achieved the lowest accuracy of 29.5% among all the prediction models for the entire test data (see Table 4.10). Moreover, models using pre-trained Word2vec, ELMo, and BERT models outperformed the baseline methods by 0.6 points (difference of 44.2% for **Word2vec (Wiki)+SVM** and 43.6% for **WordCharGram+SVM**) to 23.5 point (difference of 53.0% for **BERT (SNS)+fine-tuning** and 29.5% for **Frequency**). For sentences without emotive phrases, the proposed methods successfully achieved higher accuracy than the conventional methods by 1.6 point (difference of 38.8% for **Word2vec (Wiki)+SVM** and 37.2% for **WordCharGram+SVM**) to 29.6 point (difference of 48.4% for **BERT (SNS)+fine-tuning** and 18.8% for **Frequency**) For sentences that contain emotive expressions, the proposed models achieved better results than the baseline methods for eight emotion types except for ANGER and SHYNESS.

There were two reasons why the results of **Frequency** scored highest for inputs that contained ANGER and SHYNESS expressions. The first reason was that sentences containing expressions of ANGER or SHYNESS co-occurred with a smaller number of unique emoticons (see Figure 4.13, bars in red). Moreover, the emoticons for around 70% of sentences in the dataset were concentrated in the top 25 frequently used emoticons in the training data which is higher than the sentences containing expressions of other emotion types or sentences without emotive expressions. The second reason was that some similar sentences containing expressions of ANGER or SHYNESS co-occurred with emoticons of different emotion types which made the prediction result less adequate for the machine learning-based models. For example, Table 4.12 shows input sentences that contained the word "怒る" (*okoru*, meaning "angry"). As can be seen in Table 4.12, emoticons expressing ANGER, SADNESS, and SHYNESS are present in the training data while emoticons expressing SHYNESS, AMAZEMENT, and JOY exist in the test data. For the reference, the number of sentences that contained the word "怒る" (*okoru*, meaning "angry") in the training data was 357, and this word co-occurred with emoticons mostly expressing JOY (30.5%), SHYNESS (27.7%), and SADNESS (23.5%). The number of inputs that contained this word in the test data was 259, and it co-occurred

Table 4.11: *P* value of Wilcoxson signed-rank test between each method. "*" for values higher than 0.05. Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

Method	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1)	-										
(2)	0.00	-									
(3)	0.00	0.06*	-								
(4)	0.00	2.47e−14	1.72e−12	-							
(5)	0.00	2.21e−46	1.93e−44	1.37e−22	-						
(6)	0.00	1.18e−109	4.07e−107	2.08e−71	1.29e−23	-					
(7)	0.00	0.00	0.00	0.00	2.04e−269	1.90e−144	-				
(8)	0.00	0.00	0.00	0.00	1.31e−283	1.46e−181	1.08e−8	-			
(9)	0.00	7.45e−299	1.56e−296	2.83e−271	2.24e−142	2.39e−71	2.59e−5	4.93e−20	-		
(10)	0.00	0.00	0.00	0.00	3.08e−267	6.66e−172	6.50e−22	2.44e−10	2.12e−41	-	
(11)	0.00	0.00	0.00	0.00	0.00	0.00	5.85e−225	1.57e−183	1.10e−237	1.22e−86	-

with emoticons mostly expressing JOY (32.8%), SADNESS (25.5%), and DISLIKE (13.1%) showing that the distribution of selected emoticons is slightly different from the training data. To provide another example, Table 4.13 shows sentences that contain a word "恥ずかしい" (*hazukashii*, meaning "embarrassed"). It can be also observed that emoticons expressing DISLIKE and SHYNESS are found in the training data while emoticons expressing SADNESS and SHYNESS are present in the test data. For the reference, the number of sentences that contain the word "恥ずかしい" (*hazukashii*, meaning "embarrassed") in the training data was 211, and it co-occurred with emoticons mostly expressing SHYNESS (43.6%), JOY (29.4%), SADNESS (11.8%), and DISLIKE (9.5%). The number of inputs that contained this word in the test data was 134, and it co-occurred with emoticons mostly expressing JOY (39.6%), SADNESS (21.6%), and SHYNESS (14.5%), showing again that the distribution of the chosen emoticons is different from the training data. These two problems occurring together may lead to higher results for **Frequency** when compared to other machine learning methods for inputs containing ANGER or SHYNESS expressions.

4.5.2 (B) MAP

The results of (B) MAP (Mean Average Precision) evaluation are shown in Table 4.14. The results were also divided according to the emotion types of emotive expression in input sentences to investigate whether there is a difference in the accuracy between sentences containing an emotive expression of a certain emotion category. Overall, **BERT (SNS)+fine-tuning** achieved the best results of 0.54 point for polarity and 0.32 point for emotion categorization. This indicates that emoticons agreeing on polarity/emotion type with the correct emoticon are recommended at the top of the list especially when using this prediction model. The same tendencies were confirmed for both MAP_polarity and MAP_emotion among sentences containing different emotion type of emotive expressions and NONE type sentences. It was experimentally confirmed that the proposed methods (5)-(11) outperformed the baseline methods (1)-(4). **BERT+fine-tuning** (see (10) and (11) in Table 4.14) achieved the highest scores for sentences containing an emotive expression of each emotion type and for sentences without any emotive

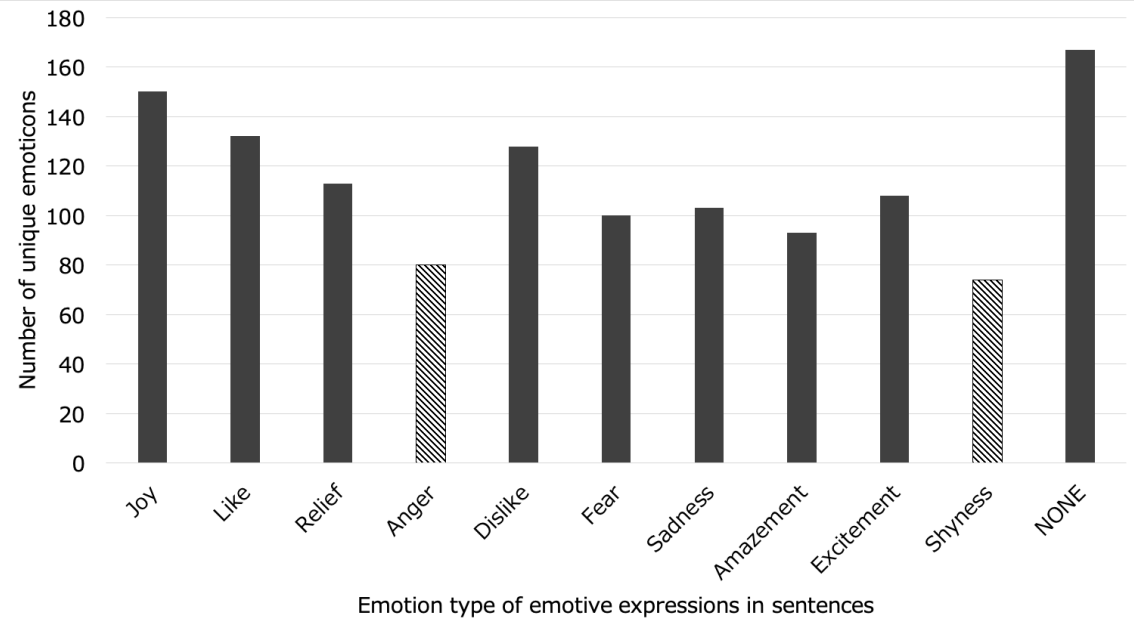


Figure 4.13: Number of unique emoticons that appeared with the emotive expression in sentences in the training data [6]. The top two lowest bars (ANGER and SHYNESS) are emphasized with stripes to mirror the discussion written in 4.5.1.

Table 4.12: Examples of sentences containing the emotive word, "怒る" (*okoru*, "angry").

Data	Sentence	Emoticon	Emotion type
Training	絶対**怒**ってるよね (He/She must be **angry**)	(｀ε ´)	ANGER
Training	ドクター**怒**っちゃうのね (The doctor gets **angry**)	(T_T)	SADNESS
Training	顔が**怒**ってらっしゃる (The face is filled with **anger**)	(^_^;)	SHYNESS
Test	めっちゃ**怒**ってるやーん (You are so **angry**)	(⌒-⌒;)	SHYNESS
Test	顔が**怒**ってるよ (Your face seems **angry**)	(; ゜ 0 ゜)	AMAZEMENT
Test	**怒**ってるように見えました (The person looked **angry**)	(^^)	SHYNESS

Table 4.13: Examples of sentences containing the emotive word, "恥ずかしい" (*hazukashii*, "embarrassed").

Data	Sentence	Emoticon	Emotion type
Training	そう言われると**恥ずかしい** (I feel **embarrassed** if you say so)	(´Д｀；)	DISLIKE
Training	見られてたとは**恥ずかしい** (You were looking at me, that's **embarrassing**)	(^_^;)	SHYNESS
Test	みじかな人に見られるとめっちゃ**恥ずかしい**です (I feel very **embarrassed** if someone I know sees me)	ヽ(；▽；)ノ	SADNESS
Test	ずっと見られると**恥ずかしい** (If someone sees me for long, I feel **embarrassed**)	(T_T)	SADNESS
Test	手相見られてたなんて何だか**恥ずかしい**です (It's kind of **embarrassing** to have my palm lines being looked at)	(*/ɪ/)	SHYNESS

expression.

Figures 4.14 and 4.15 show the distribution of average precision (AP) for polarity and emotion aspects of all samples in the test data. In these figures, AP close to 1 means that the prediction model predicts emoticons that are emotionally similar to the correct emoticon at the top of the list. AP close to 0 means that relevant emoticons are recommended at the bottom of the list. It can be observed from these figures that AP of test data predictions made by **Frequency** was mostly distributed between 0.20 to 0.60 for AP_polarity and 0 to 0.40 for AP_emotion (see (1) in Figures 4.14 and 4.15). AP of other baseline methods, **BoW+SVM**, **TF-IDF+SVM**, and **WordCharGram+SVM**, and the proposed methods **Word2vec (Wiki/SNS)+SVM**, was mostly distributed around 0.40 to 0.60 for AP_polarity and 0 to 0.60 for AP_emotion (see systems (2)-(6) in Figures 4.14 and 4.15). It was also confirmed that AP was higher (between 0.60 to 0.80) for **Word2vec (Wiki/SNS)+SVM** than other baseline methods which indicates that the proposed method using SVM recommends more relevant emoticons than the conventional methods. Other proposed methods, achieved AP higher than the aforementioned methods, mostly distributed between 0.60 to 1.00. There were only slight differences between **Word2vec (Wiki/SNS)+BiLSTM+Attention**, **ELMo+BiLSTM+Attention**, and **BERT (Wiki/SNS)+fine-tuning**. Concretely, APs of test samples predicted by **BERT (Wiki/SNS)+fine-tuning** were distributed between 0.60 to 1.00 which is higher than the other proposed methods (comparison of system models (7)-(9) and (10-11) in both Figures 4.14 and 4.15).

4.5.3 Comparison of the prediction models

According to the results of the recommendation accuracy (see Table 4.10) and MAP (Mean Average Precision) (see Table 4.14), I ranked the prediction models in the order shown in Table 4.15. From these results it can be concluded that:

- Fine-tuning the training data with the BERT model achieved the best results. Using a pre-trained model that is able to represent the contextual meaning of words may work better than pre-trained models like Word2vec which always provide an identical vector to a given word.
- Taking the word order and an attention mechanism into account improves the recommendation capability (clearly visible when **Word2vec (SNS/Wiki)+SVM** and **Word2vec (SNS/Wiki)+BiL-**

Table 4.14: MAP of each prediction model (point). "Average" means the average MAP of the entire samples. The highest values are highlighted in bold. Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

Method	Polarity	Average	Emotion type of emotive expression in inputs (with number of samples)										
	Emotion	MAP	JOY 1,930	LIKE 1,494	RELIEF 942	ANGER 420	DISLIKE 2,047	FEAR 812	SADNESS 947	AMAZEMENT 588	EXCITEMENT 720	SHYNESS 324	NONE 9,708
(1)	polarity	0.39	0.49	0.44	0.39	0.22	0.38	0.27	0.34	0.22	0.19	0.28	0.39
	emotion	0.22	0.29	0.25	0.23	0.12	0.19	0.16	0.27	0.15	0.14	0.19	0.23
(2)	polarity	0.47	0.50	0.50	0.50	0.42	0.44	0.43	0.45	0.44	0.43	0.45	0.48
	emotion	0.27	0.30	0.29	0.31	0.23	0.22	0.23	0.29	0.27	0.28	0.29	0.27
(3)	polarity	0.47	0.50	0.50	0.50	0.42	0.44	0.43	0.45	0.45	0.43	0.45	0.48
	emotion	0.27	0.30	0.29	0.31	0.23	0.22	0.24	0.29	0.28	0.28	0.29	0.27
(4)	polarity	0.47	0.50	0.49	0.50	0.42	0.44	0.43	0.45	0.44	0.44	0.45	0.48
	emotion	0.27	0.30	0.29	0.31	0.23	0.22	0.24	0.28	0.27	0.28	0.29	0.27
(5)	polarity	0.48	0.51	0.50	0.51	0.43	0.45	0.43	0.45	0.46	0.44	0.47	0.48
	emotion	0.28	0.31	0.30	0.31	0.23	0.22	0.24	0.28	0.28	0.28	0.29	0.28
(6)	polarity	0.49	0.52	0.51	0.52	0.44	0.46	0.44	0.46	0.45	0.45	0.47	0.49
	emotion	0.28	0.31	0.30	0.32	0.24	0.23	0.24	0.30	0.28	0.29	0.30	0.28
(7)	polarity	0.51	0.54	0.54	0.54	0.44	0.49	0.47	0.49	0.47	0.46	0.49	0.51
	emotion	0.30	0.33	0.32	0.33	0.24	0.24	0.25	0.33	0.29	0.29	0.31	0.30
(8)	polarity	0.51	0.55	0.55	0.55	0.44	0.48	0.45	0.49	0.47	**0.47**	0.49	0.51
	emotion	0.30	0.33	0.32	0.34	0.24	0.24	0.25	0.33	0.29	0.30	0.31	0.30
(9)	polarity	0.51	0.55	0.54	0.54	0.44	0.48	0.46	0.48	0.47	0.46	0.48	0.52
	emotion	0.30	0.33	0.32	0.33	0.24	0.24	0.25	0.32	0.29	0.29	0.30	0.30
(10)	polarity	0.53	**0.58**	0.56	0.57	0.45	0.50	**0.48**	**0.52**	0.48	**0.47**	**0.50**	0.53
	emotion	0.31	0.35	0.33	0.35	0.24	0.25	**0.26**	**0.36**	**0.31**	**0.32**	**0.32**	0.31
(11)	polarity	**0.54**	**0.58**	**0.57**	**0.58**	**0.46**	**0.51**	0.47	**0.52**	**0.49**	**0.47**	**0.50**	**0.54**
	emotion	**0.32**	**0.36**	**0.34**	**0.36**	**0.25**	**0.26**	**0.26**	0.35	**0.31**	0.31	**0.32**	**0.32**

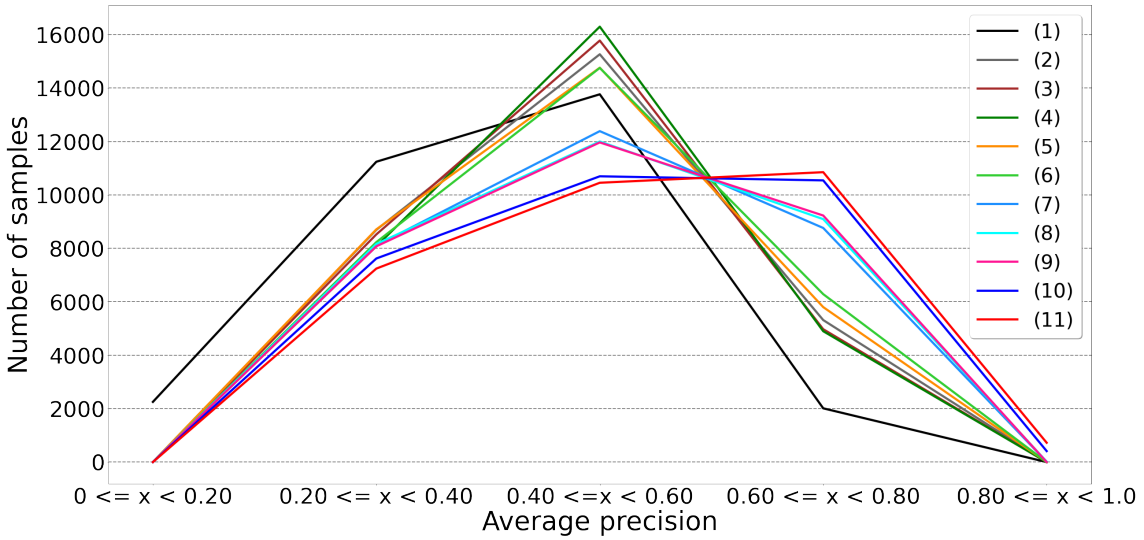


Figure 4.14: Distribution of average precision for polarity [6]. Note that sentences containing emoticons neither expressing POSITIVE nor NEGATIVE are excluded (see 4.4.1). Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

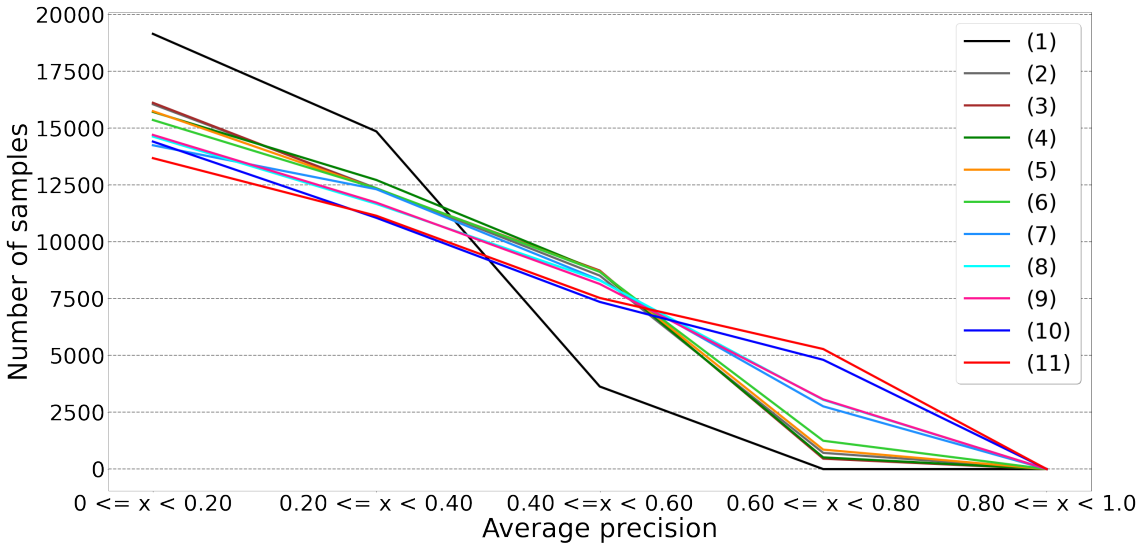


Figure 4.15: Distribution of average precision for emotion [6]. Note that sentences containing emoticons neither expressing POSITIVE nor NEGATIVE are excluded (see 4.4.1). Numbers stand for the following systems: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

Table 4.15: Ranking of the prediction models based on the results of recommendation accuracy and MAP

Method	Recommendation accuracy (%)	MAP_polarity (point)	MAP_emotion (point)	Rank
BERT (SNS)+fine-tuning	53.0	0.54	0.32	1
BERT (Wiki)+fine-tuning	50.4	0.53	0.31	2
Word2vec (SNS)+BiLSTM+Attention	48.4	0.51	0.30	3
Word2vec (Wiki)+BiLSTM+Attention	47.6	0.51	0.30	4
ELMo+BiLSTM+Attention	47.4	0.51	0.30	5
Word2vec (SNS)+SVM	45.6	0.49	0.28	6
Word2vec (Wiki)+SVM	44.2	0.49	0.28	7

STM+Attention are compared).

- Models pre-trained on the social media corpus achieved better results than the ones originating from Wikipedia articles as assumed (compare the results of the pre-trained models of **BERT+fine-tuning**, **Word2vec+BiLSTM+Attention**, and **Word2vec+SVM**). The reason may be because the way words are used in the training/test data collected from blogs and other social media is similar to the social media-based corpus used for the available pre-trained models. Moreover, more words from the test data were included in the pre-trained Word2vec (SNS) model than the model learned from Wikipedia, which I think led to better recommendation results. Considering these results, it can be expected that **ELMo+BiLSTM+Attention** may achieve a better result if the pre-trained ELMo model trained on a similar corpus was applied instead of one based on news articles.

4.5.4 Problem of recommending lower rank emoticons

BERT (SNS)+fine-tuning model achieved the best results in both recommendation accuracy and MAP (Mean Average Precision), although correct emoticons were not present in the top 25 recommendations for 47.0% (17,690 samples) of cases. From the analysis, I was able to determine two reasons. The first one is that the data size for certain emoticons was too small to enter the top 25 ranks. Figure 4.16 shows markers indicating the relation between the average ranks of emoticons recommended by each prediction model and the data size of each emoticon in the training and the test data. X-axis plots emoticon ID ranging from 1 to 167, left Y-axis represents a rank for the markers, and right Y-axis is a data size depicted as lines. As can be seen in Figure 4.16, the average rank of emoticon declines as the number of training data for the emoticon decreases. In addition, 87.0% (15,385 samples) of the 17,690 cases that **BERT (SNS)+fine-tuning** model predicted as correct emoticons from outside of the top 25 had less than 2,000 samples in the training data. I assume that the recommendation quality may improve if more samples are collected for emoticons of smaller data sizes. Although correct emoticons were not recommended at the top due to their infrequent number in the training data, the proposed methods were able to recommend emoticons of the same polarity/emotion type with the correct emoticons at the top in the recommendations (Table 4.16). Table 4.16 shows the MAP of each prediction model only focusing on emoticons that have less than 200 samples in the test data. It can be observed that the MAP results obtained by the proposed models were better than the baseline methods by 0.01 point (the difference between **Word2vec+SVM**

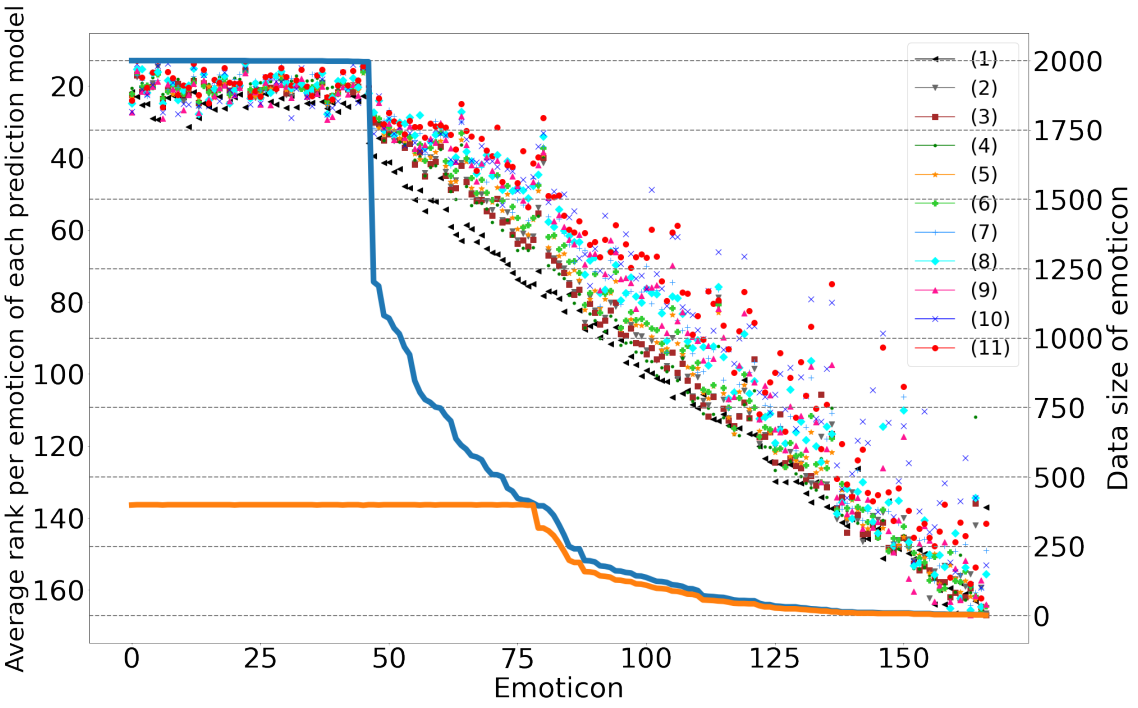


Figure 4.16: Average rank per emoticon of each prediction model (represented as markers) and data size of each emoticon in the training (blue line) and test (orange line) data [6]. Different colors mark results of following algorithms: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

Table 4.16: MAP of each prediction model for emoticons less than 200 samples in the test data. The highest values are highlighted in bold. Numbers stand for the following methods: (1) **Frequency**, (2) **BoW+SVM**, (3) **TF-IDF+SVM**, (4) **WordCharGram+SVM**, (5) **Word2vec (Wiki)+SVM**, (6) **Word2vec (SNS)+SVM**, (7) **Word2vec (Wiki)+BiLSTM+Attention**, (8) **Word2vec (SNS)+BiLSTM+Attention**, (9) **ELMo+BiLSTM+Attention**, (10) **BERT (Wiki)+fine-tuning**, (11) **BERT (SNS)+fine-tuning**.

	Method										
Type	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Polarity	0.40	0.46	0.46	0.46	0.47	0.48	0.49	0.50	0.50	0.52	**0.53**
Emotion	0.22	0.25	0.25	0.25	0.25	0.26	0.27	0.27	0.27	**0.29**	**0.29**

and **WordCharGram+SVM** to 0.12 point (the difference between **BERT (SNS)+fine-tuning** and **Frequency**). These results indicate that the proposed methods can recommend emoticons expressing similar polarity or emotion type of the correct emoticon to input sentences.

The second reason was that the tendency of emoticons used in some sentences was different from the similar sentences in the training data, thus, the prediction model could not recommend the correct emoticon within the top 25. For further analysis, I extracted 12,067 samples for which **BERT (SNS)+fine-tuning** model failed to: a) recommend correct emoticons within the top 25 and b) recommend emoticons of the same polarity/emotion type at the top of the list ($MAP < 0.50$). I investigated what kind of emoticons are inserted into similar sentences found in the training data. To do so, cosine similarity (Equation 4.12) between sentences was calculated by using the flatten layer (see Figure 4.8) extracted from **BERT (SNS)+fine-tuning** model. In Equation 4.12, $v1$ is the vector of the flatten layer for the first sentence, $v2$ is the vector of the flatten layer for the second sentence, and N is the vector size. For each of the 12,067 samples, sentences with the cosine similarity value higher than 0.90 were collected and the proportion of the emoticons of the polarity/emotion type that is the same as the emoticon in each sample was calculated. For example, there is a sentence with an emoticon "(^_^)" that belongs to the POSITIVE and JOY categories and there are 100 similar sentences. If there are 37 emoticons that express JOY and 67 emoticons that express POSITIVE among 100 sentences, then the proportion will be 37% (emotion type) and 67% (polarity type), respectively. Therefore, a higher proportion means that the emoticons of the same polarity/emotion type are also used in similar sentences in the training data. The results are shown in Table 4.17. As can be seen in Table 4.17, an average of 12.6% and 23.8% of emoticons in similar sentences had identical emotion and polarity, respectively, for 12,067 samples for which correct emoticons from the top 25 were not recommended. On the other hand, an average of 37.4% and 56.2% of emoticons in similar sentences had the same emotion type and polarity, respectively, for 19,932 samples that recommended correct emotions within the top 25. This indicates that similar sentences in the training data for 12,067 samples used emoticons differently. For example, the sentence "わからなすぎてやる気が出ない" (*wakaranasugite yarukiga denai*, meaning "I have no motivation because I don't get it at all") appeared with " (ˆ ∇ ˆ) " emoticon, expressing JOY in the test data, however, the similar sentences in the training data tend to appear with emoticons of SADNESS or DISLIKE (see Table 4.18).

Table 4.17: Average proportion of emoticons of the same polarity/emotion type used in similar sentences in the training data

Comparison	Average	
	Polarity	Emotion
12,067 samples ($rank > 25$ & $MAP < 0.50$)	23.8%	12.6%
19,932 samples ($rank <= 25$)	56.2%	37.4%

Table 4.18: Example of a sentence from training data where emoticon usage is different from the samples in the test data (the first sentence in bold is the source for finding similar sentences)

Data	Sentence	Emoticon	Emotion type
Test	わからなすぎてやる気が出ない (I have no motivation because I can't get it at all)	(^ ∇ ^)	JOY
Training	なんもやる気が出ない (I have no motivation)	(− __− ;)	DISLIKE
Training	あたしもやる気出ない (I don't have motivation either)	(¨−;)	DISLIKE
Training	やる気が起きないご (I don't feel like doing anything)	(T . T)	SADNESS
Training	掃除したいけど寒くてやる気起きません (I want to clean the room but I have no motivation because it's cold)	(; ;)	SADNESS
Training	今日は何だか何もやる気起きないんだ (I have no motivation today)	(´･_･‘)	SADNESS

$$Sim(v1, v2) = \frac{\Sigma_{i=1}^{N} v1_i v2_i}{\sqrt{\Sigma_{i=1}^{N} v1_i^2}\sqrt{\Sigma_{i=1}^{N} v2_i^2}} \tag{4.12}$$

4.6 Discussion

Overall results showed that **Deep-ERS**, especially **BERT (SNS)+fine-tuning** achieved the best emoticon recommendation performance (recommendation accuracy: 53.0%, mean average precision: 0.54 point for polarity and 0.32 point for emotion type). However, I believe that further consideration is needed to improve the recommendation performance. For example, I believe that the recommendation results improve if I incorporate the selection history of each user in **Deep-ERS**. This idea is based on my previous experiments with **EBReS-η** (see Chapter **3**). Also, I suggest employing the concept of "multimodality" [70] for **Deep-ERS**. To be more concrete, using emoticons as image data and learn the user preferences for emoticon appearance may be another idea to improve my system because I believe that users choose emoticons based on their visual information. Furthermore, the idea of "weak supervision" [71] may be useful to deal with fewer training samples for infrequently chosen emoticons. **Deep-ERS** has not been evaluated with users in the real world, therefore, my next research plan is to evaluate the system with users in the real-world environment and investigate what kind of information is necessary in order to improve **Deep-ERS**.

Aside from emoticons, emojis, stamps, and GIF images are other nonverbal information widely used

in computer-mediated communication, and users choose a suitable graphical expression from a wide variety of options in their communication. I believe that extending **Deep-ERS** to recommend other kinds of nonverbal information will broaden the support for richer online interaction. To realize this, I believe that investigation of how users make decisions of which nonverbal information to choose in their communication is necessary.

4.7 Conclusion of Chapter 4

In this chapter, I proposed a method that reorders emoticons in a suitable order based on the entire context of the input message to support users' emoticon selection. This system also aims to recommend desired emoticons to input sentences without emotive words or phrases present (Aspect (1-B)). The distinctive approach of this method was to create a prediction model that incorporates a pre-trained model (i.e. Word2vec, ELMo, and BERT) trained on large corpus in Japanese and learn patterns of sentence-emoticon pairs with machine learning method (i.e. SVM, attention-based BiLSTM, fine-tuning). Four prediction models were implemented and evaluated which one can recommend users' desired emoticons with the highest accuracy. The four prediction models were, Word2vec pre-trained model with SVM classifier, Word2vec model with attention-based BiLSTM, ELMo pre-trained model with attention-based BiLSTM, and fine-tuning with the BERT model. In addition, the influence of differences in data domains used for pre-training was examined by comparing pre-trained Word2vec and BERT models that trained on either social media corpus or Wikipedia articles. The prediction models trained on 124,359 data samples and 37,622 data samples were used for testing. The training and test data contained sentence-emoticon pairs comprising 167 unique emoticons.

From the experiment, the proposed four prediction models achieved results ranging from 44.2% to 53.0% in recommending correct emoticon within the top 25. The best method was to use fine-tuning with BERT model. These results were better than the baseline methods which ranged from 29.5% to 43.6%. Moreover, **Deep-ERS** achieved better recommendation results than the rule-based method (a method similar to Emura *et al.*'s method [31]) for sentences that contained emotive words or phrases in the input sentence (see Table 4.10). **Deep-ERS** was also discovered to be able to recommend more suitable emoticons even for sentences not containing any emotive words and phrases, an ability that cannot be achieved by the rule-based methods. I also confirmed that using a pre-trained model that trained on social media corpus achieved slightly better results (by 0.74 points to 2.55 points) than the model trained on Wikipedia articles. Some input sentences were not able to output correct emotions at the top of the list because the polarity and emotion type of emoticons used in similar sentences inside the training data were different. This means that there are patterns of how emoticons are used with messages but human-selected emoticons may differ based on other reasons (e.g. emoticon preference of users). Therefore, I believe that adding other information such as users' preferences for emoticons (e.g. proposed in Chapter **3**) could be beneficial for further improvement of the system and be more personalized to each user in the real-world applications of input methods.

Chapter 5

Conclusion

In this dissertation, I aimed to realize an emoticon recommendation system that recommends suitable emoticons to sentences inputted by users to support their emoticon selection in CMC (Computer-Mediated Communication). Firstly, I hypothesized two important aspects to realize a system that can recommend desired emoticons to a user: Aspect (1) considering the content of input sentence and Aspect (2) considering users' emoticon preference. In the previous studies, only one of the aspects was considered for recommendation; methods of Suzuki *et al.* [30] and Emura *et al.* [31] recommend emoticons based on the emotive words recognized in the input sentence (consideration of Aspect (1)) and the method currently used in most mobile operating systems which recommends emoticons selected in the past regardless of what is written in the input message (consideration of Aspect (2)). Moreover, none of these methods can recommend emoticons when emotive phrases are not present in the input message (consideration of Aspect (1)).

In Chapter **3**, I first proposed **EBReS-η** (**Emotion-prioritized History-Based Emoticon Recommendation System**) that recommends emoticons in an appropriate order to the input message by reordering emoticons based on the emotional strength of the emotion type recognized from the emotive word in the sentence and on users' selection history. With this approach, the system can recommend suitable emoticons considering both the emotion type of a message (consideration of Aspect (1)) and emoticon preference of the user who created the message (consideration of Aspect (2)). The system utilized an emoticon database that contains 59 emoticons each annotated with a strength of each 10 emotion types and a system called ML-Ask [40] that analyzes the emotion type of the entry. The proposed method was compared with the other two systems, **HiBReS** (**History-Based Emoticon Recommendation System**) and **EBReS** (**Emotion-Based Emoticon Recommendation System**). As a result, my proposed method performed best achieving 72.3% of emoticons chosen among the top five recommendations by the participants. The results for other systems ranged from 30.0% to 58.0%. Moreover, users' impression towards the proposed method was the best showing that 23 out of 30 respondents voted **EBReS-η** as the 1st performing system among the three systems. However, my proposed system failed to recommend suitable emoticons for some cases; participants chose emoticons that express different emotions from the emotion type of the emotive words recognized in the input sentences. These cases gave me the motivation to consider using machine learning techniques to discover patterns of words and emoticons and use them for emoticon recommendations.

In Chapter **4**, I proposed **Deep-Emoticon Recommendation System** (**Deep-ERS**) that can

recommend emoticons even to input messages that do not contain emotional phrases (consideration of Aspect (1)). This system was aimed to expand the capability of the recommendation from the existing systems because these systems (i.e. **EBReS-η**, Suzuki's method [30], and Emura's method [31]) can only recommend emoticons when emotional expressions appear in the entries. **Deep-ERS** can recommend emoticons in a suitable order by learning the pattern of how sentences and emoticons are used together. The system employed pre-trained models (i.e. BERT, ELMo, and Word2vec) that learned from a large Japanese corpus to grasp the context of the sentence deeper than just learning the surface of the sentences. For the evaluation experiment, I prepared prediction models that trained on 124,359 sentence-emoticon pairs and tested on 37,622 samples. As a result, the proposed method that fine-tuned with BERT achieved the best showing that 53.0% of sentences were able to retrieve the correct emoticons within the top 25. The recommendation accuracy for the other proposed methods was over 44.2% and they were better than the baseline methods which ranged from 29.5% to 43.6%. The baseline methods used were recommending frequently chosen emoticons, using bag-of-words, TF-IDF, or n-grams of words and characters as features with an SVM classifier. I also discovered that **Deep-ERS** could not recommend suitable emoticons to some sentences because the polarity and emotion type of emoticons inserted to similar sentences were different in the training data. Therefore, I believe that **Deep-ERS** may improve recommendations if other information such as users' emoticon preference is added.

For future work, I plan to evaluate **Deep-ERS** with users in the real environment to confirm its performance and also investigate what information or concepts (e.g. "multimodality", "weak supervision", etc.) are necessary for further improvement. Moreover, I plan to investigate how users decide what kind of nonverbal information to use in their communication in order to expand my system to also recommend emojis, stamps, and GIF images along with emoticons as they constantly gain popularity in computer-mediated communication.

Appendix A

Appendix for Chapter 3

A.1 Emoticon database (59 emoticon registered)

Tables A.1 and A.2 list 59 emoticons and their emotional strengths for each 10 emotion types that were used in the experiment in Chapter **3**.

A.2 Sentences used in the experiment

I prepared six groups each containing 15 sentences (five POSITIVE, five NEGATIVE, and five OTHER sentences). Sentences in each group are shown from Table A.3 to Table A.8. One group from the six groups is randomly chosen and shown to the participants during the experiment.

A.3 Factor analysis

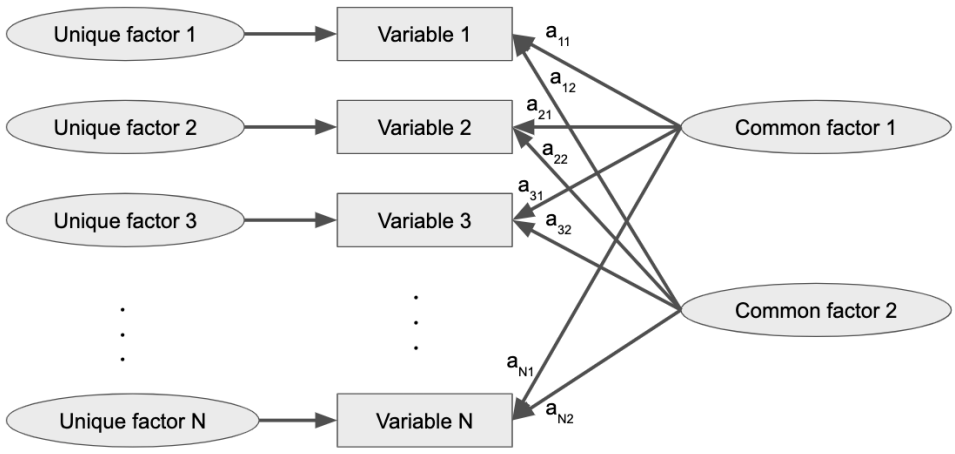


Figure A.1: Factor analysis

Factor analysis [43] is commonly used with semantic differential (SD) scale and is a method to summarize observed variables into a smaller number of unobserved factors to have a simple understanding

of users' impression (Figure A.1). In Figure A.1, Variables ("Variable 1", "Variable 2", ...,"Variable N") can be the ratings of the items given by the users. Common factors ("Common factor 1" and "Common factor 2") are the unobserved variables and are shared among two or more variables. There are usually fewer factors than variables. The unique factors ("Unique factor 1", "Unique factor 2", ..., "Unique factor N") are associated with each variable and represent the unique component of that variable. The factor model can be expressed as Equation A.1, where $x_1, x_2, ..., x_N$ derive the observed variables, $\xi_1, \xi_2, ..., \xi_k$ are the factor loadings, and $z_1, z_2, ..., z_N$ are the unique factors. The observed variables $x_1, x2, ..., x_N$ are the number of ratings for all items. $a_{11}, a_{12}, ..., a_{Nk}$ stand for factor scores which represent an association for each factor. The bigger value of the factor loading means the observed variable is strongly associated with the specific factor. Finally, factor scores and factor loadings can be used to visualize the relationship between users' impressions and the factors.

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1k} \\ a_{21} & \dots & a_{2k} \\ \vdots & \vdots & \vdots \\ a_{N1} & \dots & a_{Nk} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \tag{A.1}$$

A.4 Factor analysis result

Tables A.9, A.10, and A.11 show the detailed ratings for each of the 22 bipolar words for each system collected from 30 participants. Figure A.2 shows the correlation coefficient matrix of 22 bipolar words and eigenvalues using the aforementioned ratings. As can be seen in the red box in Figure A.2, the number of factors is decided to 4. Figure A.3 shows the result of the factor analysis when setting the number of factors to 4.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
1	(^ O ^)	0.91	0.65	0.60	0.12	0.10	0.08	0.12	0.19	0.21	0.10
2	(*^_^*)	0.87	0.78	0.65	0.10	0.08	0.06	0.09	0.10	0.16	0.25
3	(^。^)	0.78	0.60	0.56	0.11	0.10	0.08	0.12	0.20	0.20	0.17
4	(*^^*)	0.94	0.75	0.64	0.10	0.08	0.08	0.11	0.12	0.23	0.25
5	(@ ρ @)	0.29	0.25	0.42	0.22	0.38	0.13	0.17	0.23	0.20	0.32
6	(− _− ;)	0.07	0.07	0.09	0.36	0.63	0.33	0.43	0.22	0.42	0.41
7	(=^ェ^=)	0.83	0.62	0.52	0.08	0.08	0.07	0.09	0.11	0.17	0.20
8	(。-_-。)	0.25	0.32	0.34	0.23	0.38	0.18	0.37	0.10	0.17	0.45
9	(ーー ;)	0.07	0.07	0.08	0.35	0.60	0.31	0.36	0.24	0.40	0.43
10	(^◇^;)	0.26	0.16	0.18	0.18	0.37	0.30	0.18	0.49	0.50	0.55
11	(− _− #)	0.06	0.07	0.08	0.78	0.67	0.26	0.28	0.14	0.38	0.27
12	(* ` へ ´*)	0.11	0.14	0.09	0.89	0.53	0.15	0.17	0.14	0.55	0.25
13	(⌒▽⌒)	0.91	0.66	0.69	0.09	0.07	0.08	0.07	0.08	0.18	0.08
14	(# ` ∧ ´)	0.07	0.09	0.07	0.93	0.60	0.15	0.18	0.13	0.47	0.18
15	(⌒-⌒;)	0.19	0.12	0.14	0.18	0.35	0.23	0.17	0.35	0.40	0.50
16	(; O)	0.07	0.06	0.08	0.28	0.44	0.33	0.23	0.67	0.46	0.33
17	(▽)	0.53	0.39	0.51	0.12	0.19	0.09	0.11	0.23	0.23	0.21
18	(; ;)	0.10	0.09	0.09	0.14	0.38	0.49	0.86	0.10	0.28	0.27
19	(T . T)	0.12	0.06	0.10	0.18	0.43	0.45	0.79	0.15	0.25	0.24
20	(^^)	0.83	0.61	0.65	0.09	0.07	0.06	0.07	0.08	0.14	0.10
21	(^^)	0.80	0.63	0.53	0.07	0.07	0.06	0.07	0.08	0.13	0.10
22	(^_^;)	0.17	0.13	0.15	0.16	0.42	0.30	0.17	0.31	0.40	0.50
23	(^○^)	0.89	0.67	0.60	0.08	0.07	0.06	0.08	0.22	0.24	0.12
24	(´ Д ‘)	0.11	0.12	0.15	0.28	0.65	0.41	0.61	0.41	0.37	0.38
25	(^人^)	0.56	0.48	0.54	0.06	0.08	0.07	0.06	0.08	0.18	0.12
26	(T_T)	0.09	0.08	0.09	0.21	0.50	0.40	0.80	0.16	0.26	0.23
27	(; ° 0 °)	0.08	0.07	0.07	0.15	0.33	0.39	0.18	0.95	0.48	0.32
28	(´・_・‘)	0.08	0.10	0.11	0.22	0.50	0.42	0.74	0.15	0.25	0.28
29	(T ^T)	0.13	0.08	0.11	0.30	0.51	0.34	0.82	0.15	0.35	0.25
30	(・・;)	0.08	0.08	0.08	0.15	0.39	0.35	0.31	0.63	0.44	0.35

Table A.1: Emoticon database (from ID 1 to ID 30) described in Chapter **3**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
31	(#^.^#)	0.89	0.73	0.55	0.07	0.06	0.06	0.07	0.13	0.25	0.33
32	(￣ー￣)	0.24	0.19	0.33	0.27	0.37	0.15	0.16	0.13	0.21	0.18
33	(T_T)	0.08	0.08	0.12	0.20	0.51	0.47	0.87	0.19	0.25	0.29
34	(=^x^=)	0.85	0.67	0.51	0.05	0.06	0.05	0.05	0.10	0.18	0.17
35	(^_-)	0.56	0.58	0.43	0.10	0.12	0.06	0.09	0.10	0.15	0.13
36	(; ￣ェ￣)	0.06	0.08	0.08	0.30	0.62	0.34	0.28	0.46	0.45	0.42
37	(~_~;)	0.06	0.07	0.06	0.33	0.65	0.41	0.39	0.25	0.40	0.39
38	(/ _ ;)	0.06	0.11	0.08	0.14	0.47	0.49	0.88	0.20	0.24	0.33
39	(>_<)	0.28	0.25	0.18	0.23	0.47	0.50	0.58	0.38	0.45	0.57
40	(^-^)	0.87	0.67	0.63	0.08	0.09	0.07	0.07	0.10	0.20	0.09
41	(? _ ?)	0.09	0.10	0.10	0.10	0.18	0.20	0.14	0.59	0.28	0.16
42	(; ´ Д `)	0.04	0.08	0.08	0.19	0.55	0.45	0.45	0.51	0.52	0.36
43	(ｰ_ｰ;)	0.05	0.07	0.06	0.37	0.68	0.37	0.34	0.30	0.43	0.36
44	(^ー^)	0.60	0.51	0.43	0.09	0.09	0.08	0.08	0.14	0.13	0.13
45	(^O^)	0.91	0.73	0.66	0.06	0.07	0.05	0.06	0.19	0.19	0.13
46	(^ ∇ ^)	0.88	0.67	0.64	0.07	0.07	0.06	0.06	0.12	0.16	0.11
47	(^_^)	0.75	0.64	0.64	0.09	0.09	0.06	0.07	0.08	0.14	0.13
48	(o^^o)	0.88	0.81	0.68	0.06	0.06	0.05	0.05	0.08	0.21	0.25
49	(￣◇￣;)	0.08	0.09	0.08	0.27	0.54	0.38	0.28	0.67	0.48	0.40
50	(; _ ;)	0.10	0.10	0.10	0.16	0.43	0.52	0.88	0.21	0.28	0.28
51	(*^○^*)	0.96	0.81	0.64	0.05	0.05	0.05	0.06	0.19	0.24	0.20
52	(^ ◇ ^)	0.83	0.57	0.62	0.07	0.07	0.05	0.06	0.13	0.18	0.10
53	(^o^)	0.86	0.65	0.68	0.05	0.07	0.06	0.06	0.15	0.15	0.10
53	(* ` ω ´)	0.48	0.39	0.28	0.46	0.28	0.08	0.09	0.17	0.39	0.25
54	(*^o^*)	0.93	0.73	0.62	0.05	0.06	0.05	0.05	0.12	0.16	0.21
56	(・ ・?)	0.10	0.11	0.11	0.12	0.23	0.20	0.12	0.55	0.30	0.17
57	(￣(工)￣)	0.17	0.15	0.20	0.28	0.44	0.12	0.16	0.25	0.20	0.18
58	(>人<;)	0.11	0.11	0.09	0.11	0.30	0.30	0.30	0.30	0.62	0.51
59	(*☻-☻*)	0.53	0.45	0.28	0.15	0.20	0.17	0.19	0.32	0.27	0.29

Table A.2: Emoticon database (from ID 31 to ID 59) described in Chapter **3**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Sentence (original)	Translation	Emotion type
1	忘れ物が見つかって**ほっとした**	I'm so **relieved** to find my stuff	RELIEF
2	宿題が終わって**安心した**	I'm **relieved** to finish my homework	RELIEF
3	やっと**安心**することができた	I'm finally **relieved**	RELIEF
4	無事で**ほっとした**	I'm **relieved** that you're safe	RELIEF
5	これだけ広いと**のびのび**できるよね	We can **stretch out freely** with this broad space	RELIEF
6	昨日の夢は**怖かった**	My yesterday dream **scared** me	FEAR
7	将来のことが**心配だ**	I'm **worried** about my future	FEAR
8	**怯えてしまう**	I'm **afraid**	FEAR
9	先生**怖い**	My teacher is **scary**	FEAR
10	昨日変な人に後ろからついて来られて**怖かった**よ	I was **scared** because someone was following me yesterday	FEAR
11	お前の行動に**面食らった**わ	I'm **amazed** at what you have done	AMAZEMENT
12	テストの点数を見て**ショック**を受けた	I was **shocked** from my exam result	AMAZEMENT
13	断られて**ショック**を受けた	I was **shocked** because he said no	AMAZEMENT
14	突然で**びっくり**した	I was **surprised** because it was so sudden	AMAZEMENT
15	急に声をかけるから**びっくり**したよ	I got **surprised** because you suddenly talked to me	AMAZEMENT

Table A.3: Sentences used in Group 1. Emotive phrases are emphasized with bold.

ID	Sentence (original)	Translation	Emotion type
1	今日は**楽しい**夏祭り	Today, we have a **fun** summer festival	JOY
2	試験も終わったし**気楽**だ	Exams are over, I'm so **relaxed**	JOY
3	悩み事が解決して気分が**すっきり**した	I feel **better** that all my problems are behind me	JOY
4	彼女が**嬉しさ**を隠しきれずにはにかんでいる姿をみた	I could see her beaming, unable to hide her **happiness**	JOY
5	ご飯を食べれて**満足**した	I'm **satisfied** after eating food	JOY
6	(彼女・彼氏) にふられて**悲しい**	I'm **sad** that I broke up with my boyfriend	SADNESS
7	会えなくて**寂しい**な	I'm **sad** that I cannot see you	SADNESS
8	一人でクリスマスは**寂しい**よう	It's **sad** to spend Christmas alone	SADNESS
9	人目もはばからずに**泣いた**	I **cried** my eyes out	SADNESS
10	友達が留学してしまったので**悲しい**	I'm **sad** that my friend left to study abroad	SADNESS
11	起きたら８時で**焦った**	I had to **rush** because it was 8 am when I woke up	EXCITEMENT
12	終電を逃しそうになって**焦った**	I had to **rush** because I nearly missed my final train	EXCITEMENT
13	もうめっちゃ**もどかしい**んだよね	It's so **frustrating**	EXCITEMENT
14	**緊張**する	I'm **nervous**	EXCITEMENT
15	試合**緊張**するね	I'm **nervous** about today's game	EXCITEMENT

Table A.4: Sentences used in Group 2. Emotive phrases are emphasized with bold.

ID	Sentence (original)	Translation	Emotion type
1	その歌は**懐かしいね**	I'm so **relieved** to find my stuff	LIKE
2	これって**可愛い**	This is **cute**	LIKE
3	私たちの**友情**は不滅だから	Our **friendship** never ends	LIKE
4	先生に**憧れる**	I **admire** my teacher	LIKE
5	その漫画は**好き**だよ	I **like** this comic book	LIKE
6	昨日の夢は**怖**かった	My yesterday dream **scared** me	FEAR
7	将来のことが**心配**だ	I'm **worried** about my future	FEAR
8	**怯えて**しまう	I'm **afraid**	FEAR
9	先生**怖い**	My teacher is **scary**	FEAR
10	昨日変な人に後ろからついて来られて**怖**かったよ	I was **scared** because someone was following me yesterday	FEAR
11	お前の行動に**面食**らったわ	I'm **amazed** at what you have done	AMAZEMENT
12	テストの点数を見て**ショック**を受けた	I was **shocked** from my exam result	AMAZEMENT
13	断られて**ショック**を受けた	I was **shocked** because he said no	AMAZEMENT
14	突然で**びっくり**した	I was **surprised** because it was so sudden	AMAZEMENT
15	急に声をかけるから**びっくり**したよ	I got **surprised** because you suddenly talked to me	AMAZEMENT

Table A.5: Sentences used in Group 3. Emotive phrases are emphasized with bold.

ID	Sentence (original)	Translation	Emotion type
1	忘れ物が見つかって**ほっと**した	I'm so **relieved** to find my stuff	RELIEF
2	宿題が終わって**安心**した	I'm **relieved** to finish my homework	RELIEF
3	やっと**安心**することができた	I'm finally **relieved**	RELIEF
4	無事で**ほっと**した	I'm **relieved** that you're safe	RELIEF
5	これだけ広いと**のびのび**できるよね	We can **stretch out freely** with this broad space	RELIEF
6	彼は自慢話ばかりするので**嫌**だ	I **don't like** him because he brags a lot	DISLIKE
7	それは**むなしいね**	That is **vain**	DISLIKE
8	英語が話せなくて**悔しい**	It's **frustrating** that I cannot speak English	DISLIKE
9	それは**嫌**です	I **don't like** that	DISLIKE
10	あいつが**憎い**	I **hate** him	DISLIKE
11	自分の失敗に**赤面**した	I **blushed** at my failure	SHYNESS
12	それはちょっと**恥ずかしい**	That is a little **embarrassing**	SHYNESS
13	一緒にいると**照れくさい**	I feel **shy** when I'm with you	SHYNESS
14	アホなことは**恥ずかしい**	It's **embarrassing** to be an idiot	SHYNESS
15	**照れくさいねえ**	That is **embarrassing**	SHYNESS

Table A.6: Sentences used in Group 4. Emotive phrases are emphasized with bold.

ID	Sentence (original)	Translation	Emotion type
1	今日は**楽しい**夏祭り	Today, we have a **fun** summer festival	JOY
2	試験も終わったし**気楽**だ	Exams are over, I'm so **relaxed**	JOY
3	悩み事が解決して気分が**すっきり**した	I feel **better** that all my problems are behind me	JOY
4	彼女が**嬉しさ**を隠しきれずにはにかんでいる姿をみた	I could see her beaming, unable to hide her **happiness**	JOY
5	ご飯を食べれて**満足**した	I'm **satisfied** after eating food	JOY
6	弟に**怒る**	I got **angry** at my brother	ANGER
7	**腹を立て**つつ相手を立てた	I was **angry** at the person but I treated him with respect	ANGER
8	返事が遅くて**腹を立てる**	I got **angry** because his respond was late	ANGER
9	いらいらしてたので、つい**怒鳴**ってしまった	I yelled at him because I was **irritated**	ANGER
10	**怒**っているんだからね	I am **angry**	ANGER
11	自分の失敗に**赤面**した	I **blushed** at my failure	SHYNESS
12	それはちょっと**恥ずかしい**	That is a little **embarrassing**	SHYNESS
13	一緒にいると**照れくさい**	I feel **shy** when I'm with you	SHYNESS
14	アホなことは**恥ずかしい**	It's **embarrassing** to be an idiot	SHYNESS
15	**照れくさい**ねえ	That is **embarrassing**	SHYNESS

Table A.7: Sentences used in Group 5. Emotive phrases are emphasized with bold.

ID	Sentence (original)	Translation	Emotion type
1	その歌は**懐かしい**ね	I'm so **relieved** to find my stuff	LIKE
2	これって**可愛い**	This is **cute**	LIKE
3	私たちの**友情**は不滅だから	Our **friendship** never ends	LIKE
4	先生に**憧れる**	I **admire** my teacher	LIKE
5	その漫画は**好き**だよ	I **like** this comic book	LIKE
6	(彼女・彼氏) にふられて**悲しい**	I'm **sad** that I broke up with my boyfriend	SADNESS
7	会えなくて**寂しい**な	I'm **sad** that I cannot see you	SADNESS
8	一人でクリスマスは**寂しい**よう	It's **sad** to spend Christmas alone	SADNESS
9	人目もはばからずに**泣い**た	I **cried** my eyes out	SADNESS
10	友達が留学してしまったので**悲しい**	I'm **sad** that my friend left to study abroad	SADNESS
11	起きたら８時で**焦った**	I had to **rush** because it was 8 am when I woke up	EXCITEMENT
12	終電を逃しそうになって**焦った**	I had to **rush** because I nearly missed my final train	EXCITEMENT
13	もうめっちゃ**もどかしい**んだよね	It's so **frustrating**	EXCITEMENT
14	**緊張**する	I'm **nervous**	EXCITEMENT
15	試合**緊張**するね	I'm **nervous** about today's game	EXCITEMENT

Table A.8: Sentences used in Group 6. Emotive phrases are emphasized with bold.

ID	User	System	22 impression word pairs																					
			Amu	Imp	Fri	Han	Sp	Con	Nec	Lig	Rec	Fon	Nov	Sim	Int	Fre	Acc	Use	Vis	Eas	Cho	Spe	Sma	Sat
1	1	HiBReS	5	4	3	3	6	3	2	6	3	3	3	6	2	4	2	3	5	5	2	2	3	3
2	2	HiBReS	1	2	2	2	5	3	6	5	4	3	2	5	2	1	3	2	4	4	3	2	2	2
3	3	HiBReS	6	5	5	5	5	5	5	5	5	5	6	5	5	5	5	6	5	5	5	5	5	5
4	4	HiBReS	7	6	7	6	6	7	6	6	7	6	7	6	7	6	4	6	7	7	4	5	4	6
5	5	HiBReS	2	5	5	4	3	4	4	4	4	5	5	5	5	4	4	4	5	3	3	5	4	5
6	6	HiBReS	4	5	3	2	2	1	4	5	2	3	4	5	4	3	4	4	1	4	1	4	4	3
7	7	HiBReS	4	4	3	3	4	3	4	4	5	3	4	5	5	3	2	3	5	5	2	3	3	2
8	8	HiBReS	1	1	1	2	6	2	2	7	6	3	1	7	1	1	4	4	4	2	2	1	1	2
9	9	HiBReS	6	6	7	5	7	6	6	7	5	6	3	7	6	2	2	4	3	7	3	1	4	4
10	10	HiBReS	4	4	4	3	6	3	3	5	4	4	6	5	5	6	3	4	5	5	3	5	4	3
11	11	HiBReS	2	4	4	2	5	3	2	5	4	4	5	7	3	1	1	3	4	5	4	4	2	2
12	12	HiBReS	4	3	5	6	4	6	5	6	6	5	3	5	4	5	4	5	5	7	5	4	4	4
13	13	HiBReS	4	3	4	2	4	5	5	4	5	5	4	6	4	2	4	5	5	5	4	2	4	4
14	14	HiBReS	5	5	5	3	5	3	4	7	3	5	5	6	5	4	5	5	3	6	3	3	4	3
15	15	HiBReS	4	4	4	2	5	3	4	7	4	3	3	5	4	4	4	4	2	5	2	4	4	3
16	16	HiBReS	4	4	5	6	6	6	5	6	5	5	5	5	4	4	4	5	6	5	6	3	5	6
17	17	HiBReS	6	4	5	5	6	4	3	5	6	3	4	5	4	5	4	3	4	6	3	4	5	5
18	18	HiBReS	5	4	3	2	4	2	2	7	3	3	3	7	3	2	2	3	5	6	1	2	2	2
19	19	HiBReS	1	4	1	1	3	2	1	4	2	1	1	7	1	4	1	1	1	4	1	4	2	2
20	20	HiBReS	3	3	4	5	6	3	3	6	3	3	2	5	3	1	2	3	5	6	2	2	3	3
21	21	HiBReS	6	6	6	6	6	6	6	6	6	6	7	7	7	6	6	6	6	6	6	5	5	6
22	22	HiBReS	7	7	7	6	7	7	6	7	7	6	7	7	6	7	4	6	7	7	5	4	6	6
23	23	HiBReS	4	2	3	1	7	2	2	7	4	2	3	6	2	2	4	5	4	5	2	4	2	3
24	24	HiBReS	4	4	5	3	3	3	3	4	5	4	4	6	5	4	5	4	6	6	3	5	5	4
25	25	HiBReS	4	4	4	5	5	4	3	4	3	4	4	4	4	3	4	4	6	6	5	3	4	4
26	26	HiBReS	4	4	1	1	4	1	1	4	3	1	4	4	4	4	4	1	1	4	1	4	4	1
27	27	HiBReS	4	3	3	3	4	3	3	7	4	4	3	5	4	3	3	4	4	5	2	2	4	3
28	28	HiBReS	5	5	5	6	3	5	5	2	4	5	3	5	4	2	5	5	5	6	7	1	5	5
29	29	HiBReS	4	4	3	3	3	3	3	3	3	4	3	4	3	2	4	4	3	4	2	3	4	4
30	30	HiBReS	5	5	5	5	4	5	5	5	6	5	6	6	5	5	4	5	6	6	5	5	5	5

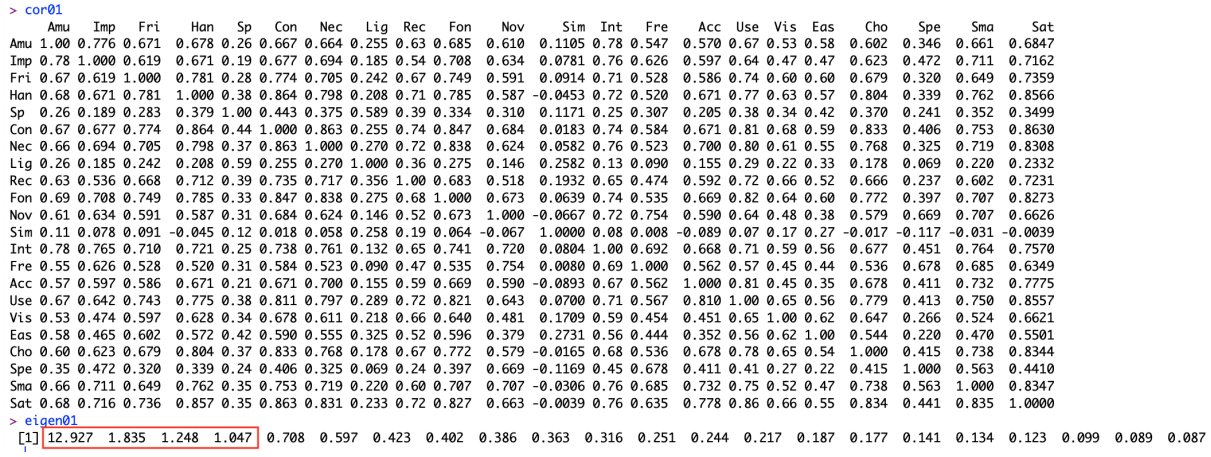
Table A.9: Ratings of HiBReS system evaluated by 30 participants.

ID	User	System	22 impression word pairs																					
			Amu	Imp	Fri	Han	Sp	Con	Nec	Lig	Rec	Fon	Nov	Sim	Int	Fre	Acc	Use	Vis	Eas	Cho	Spe	Sma	Sat
31	1	EBReS	6	6	6	6	4	6	6	6	6	6	4	6	6	4	5	6	6	6	6	2	6	7
32	2	EBReS	5	5	6	6	5	6	4	3	4	5	5	5	5	5	6	6	4	5	3	4	5	4
33	3	EBReS	5	3	4	5	5	4	4	5	6	4	5	5	3	3	4	5	4	5	5	3	4	4
34	4	EBReS	7	6	7	6	6	7	6	6	7	7	6	6	6	6	6	7	7	6	6	5	6	6
35	5	EBReS	5	5	5	5	5	5	4	5	5	5	5	4	5	5	5	5	5	5	6	4	4	5
36	6	EBReS	5	4	5	6	5	4	6	6	5	5	6	5	4	3	6	5	2	5	2	3	4	5
37	7	EBReS	5	5	6	6	4	5	5	4	6	6	4	4	5	4	4	5	6	7	7	4	4	5
38	8	EBReS	4	5	6	6	7	6	6	7	6	5	5	4	5	5	6	6	5	6	6	5	5	5
39	9	EBReS	6	6	7	7	3	7	7	4	7	6	6	7	7	7	6	7	6	7	7	2	6	7
40	10	EBReS	4	4	5	5	7	5	5	6	5	5	7	4	4	6	4	6	7	6	5	5	6	5
41	11	EBReS	5	5	3	2	5	2	4	4	5	5	5	6	5	6	3	4	2	5	3	4	4	2
42	12	EBReS	5	6	6	7	7	6	6	6	7	4	4	6	6	5	5	7	6	6	7	4	6	6
43	13	EBReS	4	4	5	5	3	5	5	3	5	5	4	5	5	3	5	5	5	5	5	4	4	5
44	14	EBReS	5	6	6	6	7	6	6	7	6	6	6	6	6	6	5	6	6	6	6	3	6	6
45	15	EBReS	3	3	4	3	7	4	3	7	3	4	3	4	3	2	4	4	3	6	4	3	4	3
46	16	EBReS	4	6	6	7	6	7	6	6	5	7	6	5	4	5	6	7	6	7	7	5	6	7
47	17	EBReS	6	7	6	7	4	6	6	5	6	7	6	5	6	6	7	7	6	6	6	5	6	7
48	18	EBReS	6	6	6	6	6	6	6	6	6	6	3	6	6	5	6	6	6	6	6	4	6	6
49	19	EBReS	3	3	3	3	7	4	4	2	4	4	3	5	5	5	4	5	7	7	3	2	3	4
50	20	EBReS	4	5	5	5	6	5	5	6	6	5	4	6	5	5	5	5	6	6	6	2	5	5
51	21	EBReS	5	6	6	7	7	6	6	6	7	6	6	6	6	6	6	6	6	6	6	6	6	6
52	22	EBReS	5	4	6	5	7	6	5	6	7	6	7	7	6	7	5	5	6	7	7	6	6	6
53	23	EBReS	6	6	4	6	7	5	5	7	6	6	4	5	5	3	6	6	3	6	5	3	4	6
54	24	EBReS	7	7	7	7	7	7	7	6	7	7	7	7	7	7	5	7	7	7	7	7	7	7
55	25	EBReS	6	6	6	5	6	6	6	6	6	6	5	6	5	4	6	6	6	6	6	4	6	5
56	26	EBReS	4	4	5	7	6	5	5	4	5	5	4	5	5	3	4	5	5	5	5	6	7	6
57	27	EBReS	6	5	6	6	7	6	5	7	7	6	6	5	6	6	5	5	7	7	5	4	5	6
58	28	EBReS	6	3	7	6	7	6	5	6	7	5	4	7	5	3	4	5	5	7	6	2	5	5
59	29	EBReS	4	4	5	5	4	5	4	5	4	5	4	5	4	4	5	4	4	5	5	4	4	4
60	30	EBReS	5	5	6	6	5	5	5	5	5	6	5	5	5	5	5	5	5	6	6	5	5	5

Table A.10: Ratings of EBReS system evaluated by 30 participants.

ID	User	System	22 impression word pairs																				
			Amu	Imp	Fri	Han	Sp	Con	Nec	Lig	Rec	Fon	Nov	Sim	Int	Fre	Acc	Use	Vis	Eas	Cho	Spe	Sma	Sat
91	1	EBReS-η	7	7	1	7	7	6	7	7	7	7	4	7	7	4	6	6	7	7	7	4	7	7
92	2	EBReS-η	5	6	6	6	6	6	7	6	6	6	6	5	5	6	7	6	6	6	5	4	5	6
93	3	EBReS-η	5	6	6	5	5	5	5	5	6	6	5	5	6	5	6	5	5	6	5	5	5	5
94	4	EBReS-η	7	6	7	7	6	6	6	6	7	6	6	6	6	6	6	7	7	7	7	6	7	7
95	5	EBReS-η	5	5	5	5	6	5	5	5	5	5	5	5	5	5	5	5	5	5	6	4	5	5
96	6	EBReS-η	6	5	6	6	6	6	5	4	5	6	5	4	5	5	5	6	6	5	7	4	4	6
97	7	EBReS-η	5	6	6	6	4	6	5	4	6	5	5	6	5	5	6	5	5	6	6	4	6	5
98	8	EBReS-η	5	5	5	6	7	6	6	7	6	6	6	3	5	5	6	6	6	6	7	6	6	6
99	9	EBReS-η	7	7	7	7	7	7	7	7	7	7	7	6	7	7	7	7	7	7	7	6	7	7
100	10	EBReS-η	4	5	5	5	7	5	5	6	5	5	6	4	5	7	5	5	5	6	6	6	6	6
101	11	EBReS-η	5	5	5	6	5	5	4	5	5	5	5	4	5	5	4	4	4	5	5	4	5	5
102	12	EBReS-η	6	4	5	6	6	6	6	5	6	4	6	5	6	4	6	6	6	6	7	3	6	6
103	13	EBReS-η	4	4	5	5	4	5	5	4	5	5	4	5	5	4	6	6	4	5	5	4	4	5
104	14	EBReS-η	5	5	6	6	6	6	6	6	6	6	6	6	6	6	5	5	5	6	6	3	6	6
105	15	EBReS-η	4	4	3	5	4	4	4	5	4	3	3	5	4	4	3	4	4	5	3	4	4	4
106	16	EBReS-η	4	6	6	7	6	7	6	6	5	7	6	5	4	5	6	7	6	7	7	5	7	7
107	17	EBReS-η	4	5	5	5	7	6	5	6	6	5	5	5	6	6	5	6	6	7	7	5	6	6
108	18	EBReS-η	6	6	6	6	4	5	4	6	6	5	3	6	5	3	5	5	5	6	7	3	5	5
109	19	EBReS-η	6	6	6	6	3	5	6	4	6	6	5	4	6	5	5	5	6	4	6	4	6	6
110	20	EBReS-η	3	3	4	5	6	6	4	6	6	5	6	3	4	3	3	3	5	5	1	3	4	3
111	21	EBReS-η	6	6	7	7	7	7	6	6	6	6	6	6	6	6	6	6	6	6	6	5	5	6
112	22	EBReS-η	7	7	7	7	7	7	7	7	7	7	7	7	6	7	6	7	7	7	7	7	7	7
113	23	EBReS-η	4	4	5	5	7	6	6	7	6	6	4	5	4	4	5	5	4	5	5	2	5	6
114	24	EBReS-η	4	3	5	5	3	4	3	3	3	5	5	5	5	5	5	5	4	5	3	4	5	4
115	25	EBReS-η	6	6	6	7	6	7	7	6	7	7	6	6	6	4	6	7	7	7	7	4	6	6
116	26	EBReS-η	6	7	4	6	6	6	5	4	4	5	7	2	6	7	4	4	5	6	6	7	6	5
117	27	EBReS-η	6	5	6	6	5	6	6	4	7	6	6	6	6	6	6	6	6	6	6	5	6	6
118	28	EBReS-η	5	5	6	5	5	5	5	4	4	4	4	4	5	4	4	4	5	6	5	4	4	5
119	29	EBReS-η	5	5	5	6	4	4	4	5	5	4	4	4	5	4	5	4	5	5	6	3	4	4
120	30	EBReS-η	5	5	5	6	6	5	5	5	6	6	5	3	5	5	6	6	5	6	6	5	5	6

Table A.11: Ratings of EBReS-η system evaluated by 30 participants.



```
> cor01
     Amu   Imp   Fri    Han   Sp   Con   Nec   Lig  Rec   Fon    Nov     Sim  Int   Fre    Acc  Use  Vis  Eas    Cho    Spe    Sma     Sat
Amu 1.00 0.776 0.671  0.678 0.26 0.667 0.664 0.255 0.63 0.685  0.610  0.1105 0.78 0.547  0.570 0.67 0.53 0.58  0.602  0.346  0.661  0.6847
Imp 0.78 1.000 0.619  0.671 0.19 0.677 0.694 0.185 0.54 0.708  0.634  0.0781 0.76 0.626  0.597 0.64 0.47 0.47  0.623  0.472  0.711  0.7162
Fri 0.67 0.619 1.000  0.781 0.28 0.774 0.705 0.242 0.67 0.749  0.591  0.0914 0.71 0.528  0.586 0.74 0.60 0.60  0.679  0.320  0.649  0.7359
Han 0.68 0.671 0.781  1.000 0.38 0.864 0.798 0.208 0.71 0.785  0.587 -0.0453 0.72 0.520  0.671 0.77 0.63 0.57  0.804  0.339  0.762  0.8566
Sp  0.26 0.189 0.283  0.379 1.00 0.443 0.375 0.589 0.39 0.334  0.310  0.1171 0.25 0.307  0.205 0.38 0.34 0.42  0.370  0.241  0.352  0.3499
Con 0.67 0.677 0.774  0.864 0.44 1.000 0.863 0.255 0.74 0.847  0.684  0.0183 0.74 0.584  0.671 0.81 0.68 0.59  0.833  0.406  0.753  0.8630
Nec 0.66 0.694 0.705  0.798 0.37 0.863 1.000 0.270 0.72 0.838  0.624  0.0582 0.76 0.523  0.700 0.80 0.61 0.55  0.768  0.325  0.719  0.8308
Lig 0.26 0.185 0.242  0.208 0.59 0.255 0.270 1.000 0.36 0.275  0.146  0.2582 0.13 0.090  0.155 0.29 0.22 0.33  0.178  0.069  0.220  0.2332
Rec 0.63 0.536 0.668  0.712 0.39 0.735 0.717 0.356 1.00 0.683  0.518  0.1932 0.65 0.474  0.592 0.72 0.66 0.52  0.666  0.237  0.602  0.7231
Fon 0.69 0.708 0.749  0.785 0.33 0.847 0.838 0.275 0.68 1.000  0.673  0.0639 0.74 0.535  0.669 0.82 0.64 0.60  0.772  0.397  0.707  0.8273
Nov 0.61 0.634 0.591  0.587 0.31 0.684 0.624 0.146 0.52 0.673  1.000 -0.0667 0.72 0.754  0.590 0.64 0.48 0.38  0.579  0.669  0.707  0.6626
Sim 0.11 0.078 0.091 -0.045 0.12 0.018 0.058 0.258 0.19 0.064 -0.067  1.0000 0.08 0.008 -0.089 0.07 0.17 0.27 -0.017 -0.117 -0.031 -0.0039
Int 0.78 0.765 0.710  0.721 0.25 0.738 0.761 0.132 0.65 0.741  0.720  0.0804 1.00 0.692  0.668 0.71 0.59 0.56  0.677  0.451  0.764  0.7570
Fre 0.55 0.626 0.528  0.520 0.31 0.584 0.523 0.090 0.47 0.535  0.754  0.0080 0.69 1.000  0.562 0.57 0.45 0.44  0.536  0.678  0.685  0.6349
Acc 0.57 0.597 0.586  0.671 0.21 0.671 0.700 0.155 0.59 0.669  0.590 -0.0893 0.67 0.562  1.000 0.81 0.45 0.35  0.678  0.411  0.732  0.7775
Use 0.67 0.642 0.743  0.775 0.38 0.811 0.797 0.289 0.72 0.821  0.643  0.0700 0.71 0.567  0.810 1.00 0.65 0.56  0.779  0.413  0.750  0.8557
Vis 0.53 0.474 0.597  0.628 0.34 0.678 0.611 0.218 0.66 0.640  0.481  0.1709 0.59 0.454  0.451 0.65 1.00 0.62  0.647  0.266  0.524  0.6621
Eas 0.58 0.465 0.602  0.572 0.42 0.590 0.555 0.325 0.52 0.596  0.379  0.2731 0.56 0.444  0.352 0.56 0.62 1.00  0.544  0.220  0.470  0.5501
Cho 0.60 0.623 0.679  0.804 0.37 0.833 0.768 0.178 0.67 0.772  0.579 -0.0165 0.68 0.536  0.678 0.78 0.65 0.54  1.000  0.415  0.738  0.8344
Spe 0.35 0.472 0.320  0.339 0.24 0.406 0.325 0.069 0.24 0.397  0.669 -0.1169 0.45 0.678  0.411 0.41 0.27 0.22  0.415  1.000  0.563  0.4410
Sma 0.66 0.711 0.649  0.762 0.35 0.753 0.719 0.220 0.60 0.707  0.707 -0.0306 0.76 0.685  0.732 0.75 0.52 0.47  0.738  0.563  1.000  0.8347
Sat 0.68 0.716 0.736  0.857 0.35 0.863 0.831 0.233 0.72 0.827  0.663 -0.0039 0.76 0.635  0.778 0.86 0.66 0.55  0.834  0.441  0.835  1.0000
> eigen01
 [1] 12.927  1.835  1.248  1.047 0.708 0.597 0.423 0.402 0.386 0.363 0.316 0.251 0.244 0.217 0.187 0.177 0.141 0.134 0.123 0.099 0.089 0.087
```

Figure A.2: Correlation of 22 bipolar words and eigenvalues. Four values in the red box are larger than 1 which is a determination of the number of factors.

```
> fa <- factanal(data2, factors = 4, scores = "Bartlett", rotation = "promax")
> fa

Call:
factanal(x = data2, factors = 4, scores = "Bartlett", rotation = "promax")

Uniquenesses:
 Amu  Imp  Fri  Han   Sp  Con  Nec  Lig  Rec  Fon  Nov  Sim  Int  Fre  Acc  Use  Vis  Eas  Cho  Spe  Sma  Sat
0.25 0.28 0.31 0.17 0.26 0.12 0.19 0.54 0.34 0.20 0.25 0.75 0.17 0.24 0.33 0.20 0.46 0.44 0.22 0.31 0.21 0.10

Loadings:
    Factor1 Factor2 Factor3 Factor4
Amu  0.356   0.178   0.518  -0.140
Imp  0.388   0.320   0.337  -0.189
Fri  0.716           0.212
Han  0.991
Sp           0.203           0.814
Con  0.924                   0.104
Nec  0.907
Lig                  0.230   0.585
Rec  0.722  -0.114   0.177   0.109
Fon  0.842
Nov  0.202   0.686
Sim -0.248  -0.135   0.548   0.127
Int  0.431   0.321   0.364  -0.165
Fre          0.811
Acc  0.833   0.144  -0.190  -0.108
Use  0.891
Vis  0.623           0.163   0.113
Eas  0.328           0.426   0.191
Cho  0.972          -0.165
Spe -0.142   0.951  -0.169   0.166
Sma  0.616   0.386
Sat  0.971          -0.113

               Factor1 Factor2 Factor3 Factor4
SS loadings       9.24    2.54   1.290   1.233
Proportion Var    0.42    0.12   0.059   0.056
Cumulative Var    0.42    0.54   0.594   0.650

Factor Correlations:
        Factor1 Factor2 Factor3 Factor4
Factor1   1.000 -0.6634 -0.2639   0.610
Factor2  -0.663  1.0000  0.0166  -0.362
Factor3  -0.264  0.0166  1.0000  -0.300
Factor4   0.610 -0.3620 -0.3001   1.000

Test of the hypothesis that 4 factors are sufficient.
The chi square statistic is 219 on 149 degrees of freedom.
The p-value is 0.00017
```

Figure A.3: Factor analysis results when the number of factors is 4

Appendix B

Appendix for Chapter 4

B.1 Emoticon database (299 emoticon registered)

Table B.1 to Table B.8 list 241 emoticons and their emotional strengths for each 10 emotion types starting from ID 60 to ID 299. These emoticons were used in the experiment in Chapter **4**. 59 emoticons shown in Appendix A.1 were also used in experiments described in Chapter **4**.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
60	+.(´^ω^‘).+	0.78	0.50	0.48	0.20	0.20	0.26	0.26	0.20	0.32	0.24
61	(‘皿´)	0.20	0.20	0.20	0.88	0.60	0.22	0.22	0.22	0.40	0.20
62	(O_O)	0.22	0.20	0.20	0.24	0.30	0.36	0.30	0.60	0.24	0.20
63	(^∀^)	0.80	0.52	0.42	0.20	0.20	0.20	0.20	0.20	0.28	0.20
64	。゜゜(つД≦。)°゜。	0.20	0.24	0.20	0.20	0.38	0.34	0.88	0.20	0.50	0.30
65	(;゜д゜)	0.20	0.20	0.20	0.20	0.30	0.38	0.22	0.76	0.34	0.26
66	(´д゜`ll)	0.20	0.20	0.20	0.22	0.60	0.40	0.48	0.30	0.22	0.28
67	σ(´～`*)	0.26	0.30	0.56	0.20	0.26	0.24	0.22	0.28	0.20	0.28
68	・゜゜・.(/□＼*).・゜゜・	0.36	0.24	0.20	0.20	0.32	0.28	0.64	0.20	0.60	0.46
69	(´Д`;)	0.20	0.20	0.20	0.30	0.64	0.32	0.42	0.26	0.20	0.26
70	。・゜゜・(>_<)・゜゜・。	0.26	0.24	0.20	0.20	0.36	0.30	0.74	0.20	0.40	0.24
71	(;一_一)	0.20	0.20	0.22	0.28	0.62	0.32	0.24	0.20	0.20	0.26
72	Σ(|||▽|||)	0.32	0.28	0.24	0.22	0.28	0.20	0.30	0.56	0.50	0.50
73	(゜,_・・゜)	0.24	0.26	0.22	0.20	0.22	0.22	0.20	0.38	0.28	0.20
74	☆(≧∀≦*)ﾉ	1.00	0.76	0.28	0.20	0.20	0.20	0.20	0.20	0.74	0.24
75	(‘^ω^)=3	0.64	0.48	0.60	0.20	0.24	0.20	0.20	0.20	0.26	0.20
76	(≡ε≡ ;)	0.24	0.28	0.32	0.26	0.48	0.20	0.20	0.22	0.24	0.48
77	Σ(゜Д゜)	0.20	0.20	0.20	0.24	0.28	0.20	0.20	0.78	0.40	0.28
78	(≧人≦)	0.34	0.26	0.20	0.20	0.24	0.28	0.40	0.20	0.42	0.60
79	(„`・ω´・)	0.38	0.36	0.24	0.26	0.20	0.20	0.20	0.20	0.54	0.22
80	(*゜ω`人)	0.54	0.56	0.36	0.20	0.20	0.20	0.20	0.20	0.30	0.22
81	ヽ(`⌒´)ﾉ	0.20	0.20	0.20	0.80	0.62	0.20	0.20	0.22	0.38	0.20
82	(`A´)	0.20	0.20	0.20	0.80	0.70	0.20	0.22	0.22	0.38	0.20
83	(*`^´)=3	0.22	0.20	0.20	0.90	0.54	0.20	0.24	0.20	0.54	0.20
84	\(//∇//)\	0.78	0.72	0.36	0.20	0.20	0.20	0.20	0.20	0.50	0.70
85	(つД‘)ノ	0.20	0.20	0.20	0.20	0.38	0.40	0.74	0.20	0.32	0.26
86	ヽ(。>д<)p	0.26	0.20	0.22	0.22	0.46	0.38	0.62	0.20	0.46	0.34
87	! (◎_◎;)	0.20	0.20	0.20	0.20	0.20	0.26	0.22	0.80	0.42	0.20
88	。゜゜(´□`。)°゜。	0.20	0.20	0.20	0.20	0.42	0.34	0.84	0.20	0.42	0.22
89	(●>ω<)ﾉ゛	0.72	0.62	0.34	0.20	0.20	0.20	0.20	0.20	0.42	0.28
90	(゜Д゜#)	0.20	0.20	0.22	0.66	0.30	0.22	0.20	0.46	0.30	0.20

Table B.1: Emoticon database (from ID 60 to ID 90) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
91	ε =(≧ m≦)	0.46	0.54	0.28	0.30	0.20	0.20	0.20	0.20	0.52	0.38
92	(≧ з≦)	0.68	0.64	0.34	0.20	0.20	0.20	0.20	0.20	0.50	0.32
93	(x д x;)	0.20	0.24	0.20	0.20	0.62	0.50	0.44	0.26	0.28	0.20
94	(ψ ` ∀ ´) ψ	0.64	0.36	0.24	0.30	0.22	0.22	0.20	0.20	0.52	0.20
95	d(˚∀˚*)	0.62	0.44	0.32	0.20	0.20	0.22	0.20	0.24	0.32	0.20
96	(|| ˚Д˚)	0.20	0.20	0.24	0.20	0.28	0.32	0.20	0.74	0.30	0.22
97	Σ (´ д ` ;)	0.20	0.20	0.22	0.20	0.36	0.30	0.30	0.66	0.20	0.26
98	o(` ω ´*)o	0.46	0.36	0.20	0.54	0.40	0.20	0.20	0.20	0.44	0.20
99	ε -(´ ∀ ` ;)	0.40	0.32	0.68	0.20	0.30	0.22	0.22	0.24	0.20	0.20
100	(* ` ω ´*)	0.50	0.36	0.20	0.54	0.24	0.20	0.20	0.20	0.46	0.22
101	。・°°・(>_<)・°°・。	0.28	0.22	0.20	0.20	0.46	0.32	0.76	0.22	0.44	0.24
102	(/-＼)	0.28	0.22	0.24	0.20	0.30	0.20	0.34	0.20	0.24	0.62
103	(*>艸<)	0.62	0.50	0.30	0.20	0.20	0.20	0.20	0.20	0.52	0.44
104	(u ω u ;A)	0.20	0.28	0.42	0.20	0.32	0.26	0.36	0.26	0.24	0.22
105	ヾ (●・v・人・v・○)ﾉ	0.58	0.68	0.50	0.20	0.20	0.20	0.20	0.20	0.30	0.20
106	(* ` 皿 ´*)ﾉ	0.36	0.28	0.26	0.84	0.34	0.20	0.20	0.20	0.54	0.26
107	。+˚ (。` Д ´。)˚+。	0.36	0.20	0.20	0.72	0.22	0.22	0.42	0.20	0.56	0.20
108	。・˚・(つД `)・˚・。	0.24	0.20	0.20	0.20	0.38	0.30	0.88	0.20	0.50	0.20
109	o(>ω<*)o	0.74	0.50	0.22	0.20	0.20	0.20	0.20	0.20	0.50	0.36
110	(● ´ Д `)ﾉﾞ	0.32	0.20	0.36	0.20	0.34	0.20	0.42	0.20	0.20	0.20
111	(((o(*˚▽˚*)o)))	0.80	0.60	0.24	0.20	0.20	0.20	0.20	0.22	0.76	0.20
112	(´・ω・lll)	0.22	0.20	0.20	0.20	0.56	0.56	0.54	0.24	0.20	0.26
113	Σ (||˚Д˚)	0.20	0.20	0.20	0.20	0.28	0.56	0.22	0.88	0.34	0.22
114	。・˚・(ﾉД ‘)・˚・。	0.20	0.20	0.20	0.20	0.42	0.32	0.86	0.20	0.40	0.20
115	(˚ д˚)	0.20	0.22	0.24	0.22	0.26	0.22	0.20	0.58	0.28	0.20
116	Σ (´˚ д˚ ‘)	0.20	0.20	0.20	0.20	0.30	0.28	0.32	0.82	0.32	0.20
117	(* ´ 艸 `)	0.66	0.54	0.32	0.20	0.20	0.20	0.20	0.20	0.26	0.24
118	(ˆ ω ˆ)	0.74	0.46	0.48	0.24	0.20	0.20	0.20	0.20	0.28	0.20
119	(ﾉ˚∀˚)ﾉ	0.76	0.42	0.34	0.20	0.20	0.20	0.20	0.24	0.46	0.20
120	o(≧ ω≦)o	0.96	0.64	0.20	0.20	0.20	0.20	0.24	0.20	0.72	0.26

Table B.2: Emoticon database (from ID 91 to ID 120) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
121	(○´艸`)	0.66	0.56	0.36	0.20	0.20	0.20	0.20	0.20	0.34	0.26
122	ヾ(●´∀`●)	0.70	0.62	0.62	0.20	0.20	0.20	0.20	0.20	0.26	0.20
123	(;¬__¬)	0.20	0.20	0.22	0.24	0.66	0.32	0.22	0.20	0.26	0.30
124	(´-‘*))))	0.34	0.32	0.70	0.20	0.20	0.20	0.20	0.20	0.22	0.24
125	(/□≦、)	0.22	0.20	0.20	0.22	0.40	0.38	0.66	0.20	0.46	0.40
126	(*´∀`*)	0.76	0.68	0.60	0.20	0.20	0.20	0.20	0.20	0.36	0.20
127	(`∀´)	0.60	0.24	0.20	0.30	0.20	0.20	0.20	0.20	0.36	0.20
128	ヾ(^▽^ゞ)	0.82	0.46	0.20	0.20	0.20	0.20	0.20	0.20	0.34	0.20
129	ヽ(‘З’)ﾉ	0.20	0.20	0.20	0.66	0.46	0.20	0.24	0.22	0.48	0.20
130	(*ﾉɪﾉ)	0.30	0.38	0.22	0.20	0.20	0.22	0.22	0.20	0.26	0.56
131	ヽ(≧∀≦)ﾉ	0.92	0.48	0.20	0.28	0.20	0.20	0.20	0.20	0.62	0.30
132	(°△°;ﾉ)ﾉ	0.20	0.20	0.20	0.22	0.38	0.44	0.20	0.72	0.26	0.20
133	(´Д‘)	0.22	0.22	0.28	0.22	0.54	0.36	0.44	0.22	0.22	0.22
134	(●´З`●)	0.42	0.46	0.56	0.22	0.26	0.20	0.24	0.20	0.22	0.22
135	((=°Д°=))ﾉ	0.28	0.28	0.20	0.38	0.28	0.22	0.20	0.32	0.28	0.24
136	(≧Д≦)	0.38	0.34	0.24	0.22	0.34	0.24	0.36	0.20	0.56	0.34
137	ΣΣ(°д°lll)!	0.20	0.20	0.20	0.20	0.46	0.56	0.20	0.84	0.42	0.26
138	。°(。ﾉω丶。)°。	0.20	0.26	0.22	0.20	0.30	0.26	0.74	0.20	0.42	0.20
139	(*б ω б*)	0.40	0.52	0.28	0.20	0.20	0.20	0.28	0.20	0.34	0.34
140	(≧ω≦)	0.80	0.52	0.32	0.20	0.20	0.20	0.20	0.20	0.60	0.28
141	(〃°A°)	0.26	0.38	0.26	0.26	0.24	0.24	0.20	0.36	0.24	0.46
142	(OvO)	0.54	0.28	0.30	0.22	0.20	0.20	0.20	0.28	0.28	0.24
143	llllll(-ω-;)llllll	0.20	0.20	0.20	0.20	0.72	0.46	0.40	0.30	0.26	0.24
144	Σ(°д°|||)	0.20	0.20	0.20	0.20	0.46	0.42	0.22	0.90	0.28	0.20
145	(°д°)	0.20	0.20	0.20	0.20	0.26	0.22	0.20	0.60	0.24	0.20
146	+°。*(*´∀`*)*。°+	0.88	0.58	0.38	0.22	0.20	0.20	0.20	0.20	0.64	0.22
147	(*≧∪≦)	0.88	0.66	0.28	0.20	0.20	0.20	0.20	0.20	0.50	0.32
148	(人ｕдｕ)	0.22	0.34	0.46	0.20	0.24	0.20	0.26	0.20	0.22	0.32
149	(〃▽〃)	0.48	0.58	0.26	0.20	0.20	0.20	0.20	0.20	0.24	0.72
150	(`Д´)	0.20	0.20	0.20	0.80	0.44	0.24	0.20	0.20	0.38	0.20

Table B.3: Emoticon database (from ID 121 to ID 150) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
151	(; _ ;)/	0.20	0.28	0.20	0.20	0.36	0.20	0.78	0.20	0.26	0.24
152	(;゜Д゜ i|!)	0.20	0.20	0.20	0.20	0.46	0.56	0.30	0.78	0.30	0.20
153	(。・｀ω・。)q	0.38	0.34	0.20	0.22	0.20	0.20	0.20	0.20	0.52	0.20
154	Σ (ll゜Д゜ ll)	0.20	0.20	0.20	0.20	0.36	0.46	0.32	0.94	0.32	0.26
155	Σ (゜д゜;)ﾉ	0.20	0.20	0.20	0.20	0.36	0.46	0.20	0.90	0.40	0.20
156	.+゜.+゜ (o(。・д・。)o).+゜.+゜	0.64	0.48	0.38	0.20	0.20	0.20	0.20	0.20	0.64	0.20
157	(*≧ д≦)	0.54	0.48	0.30	0.26	0.28	0.20	0.20	0.20	0.56	0.48
158	ヾ (*´ｴ`*)ﾉ	0.50	0.54	0.70	0.20	0.20	0.20	0.20	0.20	0.36	0.24
159	(*◎ v ◎*)	0.62	0.50	0.38	0.20	0.20	0.20	0.20	0.32	0.38	0.22
160	ヽ (=´ Д `=)ﾉ	0.32	0.28	0.38	0.24	0.50	0.30	0.42	0.20	0.22	0.24
161	(〃 ω 〃)	0.34	0.48	0.24	0.20	0.20	0.20	0.20	0.20	0.28	0.70
162	(つ ﾛ⊂)	0.20	0.20	0.20	0.20	0.28	0.28	0.56	0.20	0.22	0.38
163	(。-_-。)ﾉ	0.24	0.24	0.26	0.24	0.48	0.20	0.30	0.20	0.20	0.24
164	(' Д `◆)	0.20	0.20	0.22	0.20	0.64	0.26	0.34	0.22	0.20	0.20
165	(´ Д ` υ)	0.22	0.20	0.24	0.20	0.68	0.24	0.36	0.20	0.20	0.20
166	(ﾉД `)	0.20	0.20	0.20	0.20	0.60	0.28	0.58	0.20	0.20	0.20
167	(＃゜д゜ﾒ)	0.20	0.20	0.20	0.68	0.24	0.20	0.20	0.36	0.36	0.20
168	(*´v ｀)	0.52	0.68	0.64	0.20	0.20	0.20	0.20	0.20	0.20	0.20
169	[● ´ ; д ;]	0.26	0.20	0.20	0.20	0.34	0.36	0.82	0.20	0.30	0.22
170	(´ △ ｀)	0.24	0.20	0.28	0.24	0.38	0.26	0.40	0.20	0.20	0.20
171	(*p ω q*)	0.54	0.46	0.30	0.20	0.20	0.20	0.20	0.24	0.32	0.44
172	(｀ ε ´)	0.20	0.20	0.20	0.64	0.60	0.20	0.20	0.20	0.28	0.22
173	d(´ ω ｀*)	0.54	0.50	0.60	0.20	0.20	0.20	0.20	0.20	0.26	0.22
174	(„＃゜Д゜)	0.28	0.20	0.20	0.60	0.34	0.22	0.20	0.40	0.36	0.28
175	(・ ´ ω `・)	0.24	0.26	0.34	0.26	0.32	0.30	0.42	0.20	0.28	0.22
176	♪———O （≧∇≦） O———♪	1.00	0.62	0.20	0.20	0.20	0.20	0.20	0.20	0.96	0.20
177	d(* ｀ c_゜ ´*)b	0.62	0.40	0.20	0.32	0.20	0.20	0.20	0.20	0.62	0.24
178	＼(^o^)／	0.72	0.28	0.28	0.20	0.20	0.20	0.38	0.20	0.38	0.20
179	(*ﾉ∀゜*) σ	0.64	0.46	0.28	0.20	0.20	0.20	0.20	0.20	0.36	0.44
180	.+(´ ^ ω ^ `)+.	0.78	0.36	0.50	0.20	0.20	0.20	0.26	0.20	0.24	0.20

Table B.4: Emoticon database (from ID 151 to ID 180) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
181	(;;˚:ж:˚:;)	0.26	0.20	0.24	0.20	0.36	0.30	0.30	0.36	0.34	0.26
182	o(‘･ω ´･+o)	0.42	0.38	0.24	0.24	0.20	0.20	0.20	0.20	0.56	0.20
183	ヽ(｀Д´)ﾉ	0.20	0.20	0.20	0.80	0.54	0.24	0.26	0.20	0.38	0.20
184	(。u ω u。)	0.26	0.40	0.54	0.20	0.20	0.24	0.32	0.24	0.20	0.30
185	(◆*’v‘*)	0.56	0.50	0.46	0.20	0.20	0.20	0.20	0.20	0.30	0.28
186	.+˚ヽ(｀･ω･´)ﾉ.+˚	0.54	0.42	0.26	0.20	0.20	0.20	0.20	0.20	0.68	0.20
187	(*’ω｀*)ゞ	0.56	0.50	0.40	0.20	0.20	0.20	0.20	0.20	0.30	0.32
188	(●｀艸´)	0.36	0.34	0.20	0.38	0.28	0.20	0.20	0.24	0.54	0.20
189	(〃´皿‘)q	0.26	0.22	0.26	0.26	0.52	0.22	0.24	0.20	0.26	0.32
190	(●´∀｀●)ﾉ	0.60	0.50	0.54	0.20	0.20	0.20	0.20	0.20	0.32	0.24
191	(；д；)	0.20	0.20	0.20	0.20	0.30	0.32	0.82	0.20	0.24	0.20
192	(((((；˚Д˚)))))))	0.24	0.20	0.20	0.20	0.42	0.68	0.22	0.72	0.48	0.20
193	Σ(o˚ε˚´oﾉ)ﾉ	0.26	0.30	0.28	0.20	0.20	0.20	0.20	0.72	0.32	0.24
194	(‘･ω ´･●)	0.32	0.28	0.26	0.26	0.20	0.20	0.20	0.20	0.44	0.22
195	(｀･ω ´･+)v	0.32	0.28	0.28	0.26	0.20	0.20	0.20	0.20	0.50	0.20
196	。･˚･(ﾉ∀ ‘)･˚･。	0.70	0.42	0.30	0.20	0.20	0.20	0.32	0.20	0.56	0.22
197	ヾ(o´∀｀o)ﾉ	0.84	0.60	0.30	0.20	0.20	0.20	0.20	0.20	0.46	0.20
198	(*ﾉωﾉ)	0.34	0.54	0.38	0.20	0.20	0.20	0.24	0.20	0.24	0.56
199	(¬ω¬;A)	0.20	0.20	0.20	0.20	0.56	0.28	0.26	0.20	0.26	0.26
200	(●´∪｀)	0.38	0.44	0.68	0.20	0.20	0.20	0.20	0.20	0.20	0.20
201	(¬з¬)	0.20	0.20	0.22	0.26	0.48	0.22	0.22	0.22	0.28	0.20
202	d(･｀ω ´･d*)	0.46	0.34	0.24	0.26	0.20	0.20	0.20	0.20	0.44	0.20
203	(人´∀｀).☆.。.:*･˚	0.70	0.62	0.34	0.20	0.20	0.20	0.20	0.20	0.46	0.20
204	((*´艸｀))	0.66	0.40	0.36	0.20	0.24	0.20	0.20	0.20	0.30	0.38
205	σ(0●_0*)	0.20	0.20	0.22	0.20	0.20	0.20	0.20	0.54	0.28	0.26
206	《(;´Д‘)》	0.20	0.20	0.20	0.20	0.56	0.56	0.30	0.30	0.30	0.22
207	(。≧ω≦)	0.62	0.46	0.20	0.20	0.22	0.22	0.30	0.22	0.46	0.30
208	Ψ(｀∀´)Ψ	0.46	0.26	0.20	0.32	0.30	0.20	0.24	0.20	0.54	0.20
209	(o‿ω‿o)	0.64	0.52	0.50	0.20	0.20	0.20	0.20	0.20	0.28	0.22
210	(●｀･ω･)	0.34	0.32	0.22	0.28	0.20	0.20	0.20	0.20	0.30	0.20

Table B.5: Emoticon database (from ID 181 to ID 210) described in Chapter **4**. “SAD” for SADNESS, “AMA” for AMAZEMENT, “EXC” for EXCITEMENT, and “SHY” for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
211	il||li(つд-。)il||li	0.20	0.20	0.26	0.22	0.58	0.54	0.64	0.20	0.30	0.20
212	(´A ｀)	0.20	0.20	0.20	0.20	0.56	0.30	0.34	0.20	0.20	0.20
213	o(；>△<)o	0.20	0.20	0.20	0.20	0.52	0.52	0.42	0.26	0.44	0.24
214	σ（o д o lll)	0.20	0.24	0.20	0.20	0.46	0.44	0.40	0.58	0.30	0.28
215	(*´ɪ‘*)	0.46	0.54	0.58	0.20	0.20	0.20	0.20	0.20	0.20	0.24
216	σ (-c_-´。)	0.20	0.20	0.28	0.22	0.44	0.24	0.28	0.20	0.24	0.20
217	(つД｀)	0.20	0.20	0.20	0.20	0.46	0.30	0.52	0.22	0.26	0.20
218	(=´ω｀=)	0.48	0.46	0.60	0.20	0.20	0.20	0.26	0.20	0.20	0.30
219	･ﾟ･(つД｀)･ﾟ･	0.20	0.20	0.20	0.20	0.32	0.32	0.72	0.20	0.50	0.20
220	(=ﾟДﾟ)ﾉ	0.20	0.20	0.20	0.42	0.24	0.20	0.20	0.36	0.40	0.20
221	((((；ﾟДﾟ))))	0.20	0.20	0.20	0.20	0.34	0.60	0.22	0.58	0.36	0.22
222	Σ【*ﾟдﾟ*】	0.34	0.30	0.20	0.20	0.20	0.20	0.20	0.82	0.32	0.42
223	Σ(;ﾟдﾟ)	0.20	0.20	0.20	0.20	0.28	0.30	0.20	0.80	0.30	0.24
224	(ﾉД‘)	0.20	0.20	0.20	0.20	0.44	0.24	0.58	0.20	0.20	0.20
225	ﾟ。(*^▽^*) ﾚﾞ	0.76	0.52	0.26	0.20	0.20	0.20	0.20	0.20	0.36	0.24
226	d(ﾟДﾟ*)	0.40	0.30	0.28	0.20	0.20	0.22	0.20	0.32	0.46	0.26
227	(　｀∀´)	0.42	0.20	0.28	0.24	0.26	0.20	0.20	0.20	0.44	0.20
228	+.ﾟ(*´∀‘)bﾟ+.ﾟ	0.80	0.56	0.56	0.20	0.20	0.20	0.20	0.20	0.48	0.20
229	(σ´Д｀)σ	0.24	0.20	0.28	0.22	0.34	0.22	0.36	0.20	0.34	0.20
230	ヽ(*´∀｀*)ﾉ.+ﾟ	0.84	0.50	0.48	0.20	0.20	0.20	0.20	0.20	0.48	0.20
231	ε-(o´ω｀o)	0.34	0.46	0.74	0.20	0.20	0.20	0.20	0.20	0.20	0.20
232	σ(≧ε≦o)	0.84	0.38	0.24	0.20	0.20	0.20	0.20	0.20	0.50	0.32
233	(´；ω；｀)	0.20	0.20	0.20	0.20	0.36	0.34	0.80	0.20	0.32	0.20
234	(。v ω v。｀)	0.22	0.40	0.42	0.20	0.24	0.22	0.34	0.20	0.20	0.30
235	δ(´д｀;)	0.20	0.20	0.22	0.20	0.54	0.24	0.24	0.22	0.24	0.20
236	(´艸｀)	0.52	0.28	0.24	0.20	0.24	0.24	0.20	0.24	0.26	0.24
237	(ー д ー;)	0.20	0.20	0.20	0.26	0.52	0.26	0.28	0.20	0.22	0.22
238	(^з^)-☆	0.56	0.54	0.24	0.20	0.20	0.20	0.20	0.20	0.32	0.20
239	(ﾟ∇^*)	0.70	0.42	0.24	0.20	0.20	0.20	0.20	0.20	0.36	0.22
240	(●ﾟωﾟ●)	0.48	0.44	0.34	0.20	0.20	0.20	0.20	0.28	0.34	0.22

Table B.6: Emoticon database (from ID 211 to ID 240) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
241	(u_u)	0.20	0.28	0.38	0.20	0.26	0.20	0.34	0.20	0.20	0.24
242	[*´・ω・`]	0.20	0.20	0.20	0.20	0.32	0.26	0.56	0.20	0.28	0.22
243	(` Д ´)/	0.20	0.20	0.20	0.78	0.36	0.20	0.20	0.20	0.44	0.20
244	ヽ (; ▽ ;) ノ	0.44	0.36	0.20	0.20	0.24	0.20	0.58	0.20	0.34	0.24
245	(*ノ∀ ‘*)	0.70	0.56	0.36	0.20	0.20	0.20	0.26	0.20	0.34	0.44
246	w(゜△゜)w	0.20	0.20	0.20	0.28	0.32	0.22	0.20	0.66	0.26	0.24
247	(*´‘ω ´)/	0.32	0.38	0.20	0.50	0.36	0.20	0.20	0.20	0.46	0.20
248	(´∀ ` *)	0.56	0.46	0.58	0.20	0.20	0.20	0.20	0.20	0.20	0.20
249	(メ=д=o)	0.26	0.28	0.22	0.28	0.58	0.22	0.26	0.20	0.26	0.26
250	゜+('v‘*)+.゜	0.62	0.60	0.38	0.20	0.20	0.20	0.20	0.20	0.24	0.32
251	ヾ (゜∀゜ゞ)	0.54	0.38	0.32	0.20	0.20	0.20	0.20	0.24	0.52	0.20
252	(´・ω・‘)	0.20	0.20	0.22	0.20	0.36	0.28	0.68	0.20	0.24	0.20
253	。゜(゜´Д`゜)゜。	0.20	0.20	0.20	0.20	0.46	0.34	0.76	0.20	0.42	0.20
254	(` ・ω・ ´)	0.22	0.20	0.24	0.36	0.22	0.20	0.20	0.20	0.54	0.20
255	(` ・д・ ´)	0.20	0.20	0.20	0.50	0.24	0.20	0.20	0.20	0.52	0.20
256	(*^▽^*)	0.80	0.48	0.38	0.20	0.20	0.20	0.20	0.20	0.34	0.20
257	(*≧ m ≦ *)	0.64	0.44	0.28	0.20	0.20	0.20	0.20	0.20	0.58	0.42
258	ε =(> ε <)	0.36	0.28	0.36	0.26	0.26	0.20	0.20	0.26	0.40	0.20
259	(` ・∀ ・+ ´)	0.42	0.34	0.20	0.22	0.20	0.20	0.20	0.20	0.56	0.20
260	♪(´▽ `)	0.58	0.46	0.66	0.20	0.20	0.20	0.20	0.20	0.30	0.20
261	(」 *´∇ `)」	0.64	0.50	0.64	0.20	0.20	0.20	0.20	0.20	0.40	0.20
262	(„人゜Д゜)	0.24	0.28	0.24	0.20	0.24	0.20	0.20	0.46	0.38	0.22
263	(^ω^)	0.60	0.50	0.50	0.20	0.20	0.20	0.20	0.20	0.22	0.20
264	Σ (・□・ ;)	0.20	0.20	0.20	0.20	0.22	0.26	0.22	0.72	0.24	0.20
265	(´艸‘*)	0.60	0.46	0.40	0.22	0.22	0.20	0.20	0.20	0.24	0.50
266	(((゜Д゜)))	0.20	0.20	0.20	0.20	0.34	0.50	0.20	0.76	0.28	0.20
267	(゜Д゜)	0.20	0.20	0.20	0.20	0.30	0.32	0.22	0.68	0.26	0.20
268	(* ゜Д゜)	0.20	0.24	0.20	0.28	0.26	0.30	0.20	0.60	0.28	0.26
269	(o´▽ ` o)	0.64	0.48	0.78	0.20	0.20	0.20	0.20	0.20	0.20	0.22
270	(´‘ω ´)	0.24	0.20	0.20	0.60	0.28	0.20	0.20	0.20	0.34	0.20

Table B.7: Emoticon database (from ID 241 to ID 270) described in Chapter **4**. "SAD" for SADNESS, "AMA" for AMAZEMENT, "EXC" for EXCITEMENT, and "SHY" for SHYNESS.

ID	Emoticon	emotional strength									
		JOY	LIKE	RELIEF	ANGER	DISLIKE	FEAR	SAD	AMA	EXC	SHY
271	(● ´A ｀)	0.20	0.22	0.24	0.22	0.44	0.30	0.36	0.20	0.22	0.20
272	ヽ (´ ρ ｀)ﾉ	0.28	0.20	0.32	0.20	0.34	0.20	0.32	0.20	0.20	0.20
273	Σ (´ Д ｀　)	0.20	0.20	0.20	0.22	0.40	0.28	0.36	0.68	0.28	0.20
274	(ﾉ∇ ≦ *)	0.66	0.56	0.24	0.20	0.20	0.20	0.20	0.20	0.52	0.54
275	ヽ (´ A ｀)ﾉ	0.22	0.20	0.22	0.20	0.40	0.24	0.38	0.20	0.20	0.20
276	(。 ≧ ω≦)ﾉ	0.72	0.56	0.28	0.20	0.22	0.20	0.24	0.20	0.50	0.32
277	。 ・゜ ・(ノД ‘) ・゜ ・。	0.26	0.20	0.20	0.20	0.40	0.32	0.84	0.20	0.56	0.20
278	(゜Д゜)	0.20	0.20	0.24	0.28	0.36	0.24	0.20	0.58	0.24	0.20
279	(/∀ ＼*)	0.58	0.50	0.24	0.20	0.20	0.20	0.20	0.20	0.20	0.80
280	(*u ω u*) σ	0.48	0.56	0.44	0.20	0.20	0.20	0.20	0.20	0.28	0.52
281	(´ ⊥ ｀*)	0.38	0.48	0.58	0.20	0.20	0.20	0.20	0.20	0.20	0.28
282	(´ ＿ ｀。)	0.24	0.30	0.40	0.20	0.28	0.24	0.38	0.20	0.20	0.24
283	*・゜ ゜・*:.。..。.:*・’(*゜▽゜*)’・*:.。. .。.:*・゜ ゜・*	0.86	0.54	0.30	0.20	0.20	0.20	0.20	0.20	0.66	0.20
284	(*v ω v)	0.34	0.52	0.46	0.20	0.20	0.20	0.22	0.20	0.28	0.42
285	(#゜Д゜)	0.20	0.26	0.20	0.64	0.28	0.20	0.20	0.38	0.42	0.20
286	(≧ 艸≦ *)	0.82	0.44	0.22	0.20	0.20	0.20	0.20	0.20	0.52	0.64
287	ш (゜Д゜ш)	0.26	0.28	0.26	0.36	0.38	0.26	0.24	0.48	0.64	0.20
288	(。 ・ω・。)	0.28	0.48	0.52	0.20	0.22	0.20	0.22	0.20	0.20	0.20
289	(σ・Д・) σ	0.24	0.30	0.32	0.22	0.26	0.22	0.20	0.28	0.36	0.20
290	(o-_-o)	0.20	0.22	0.28	0.22	0.32	0.24	0.34	0.24	0.20	0.22
291	Σ d(≧ ∀ ≦ *)	0.92	0.52	0.20	0.20	0.20	0.20	0.20	0.26	0.70	0.22
292	Σ (=゜□゜= ;)	0.38	0.32	0.20	0.20	0.22	0.24	0.20	0.82	0.32	0.36
293	(｀゜Д゜ ´) ﾍﾟ	0.24	0.22	0.22	0.28	0.20	0.22	0.20	0.32	0.62	0.24
294	☆*:.。. o(≧ ▽≦)o .。.:*☆	0.98	0.52	0.20	0.20	0.20	0.20	0.20	0.20	0.84	0.28
295	ヾ (´ ε ｀;) ﾉ	0.20	0.22	0.28	0.20	0.32	0.24	0.20	0.20	0.32	0.26
296	(„゜Д゜)	0.22	0.26	0.20	0.20	0.20	0.26	0.24	0.66	0.26	0.30
297	(*´ ω ｀*)	0.58	0.70	0.50	0.20	0.20	0.20	0.20	0.20	0.30	0.30
298	(〃 ´3’ 〃) ﾍﾟ	0.56	0.54	0.24	0.20	0.20	0.20	0.20	0.20	0.36	0.54
299	ヽ (;´ Д ｀)ﾉ	0.20	0.22	0.24	0.22	0.52	0.42	0.42	0.30	0.24	0.20

Table B.8: Emoticon database (from ID 271 to ID 299) described in Chapter **4**. “SAD” for SADNESS, “AMA” for AMAZEMENT, “EXC” for EXCITEMENT, and “SHY” for SHYNESS.

Achknowledgements

I would like to thank the following people, without whom I would not have been able to complete my research and receive a Ph.D. degree.

First of all, I would like to express the deepest appreciation to my doctoral course academic adviser Professor Kenji Araki (Language Media Laboratory, Faculty of Information Science and Technology, Hokkaido University), for providing guidance and feedback, encouragement and patience, throughout my journey to the highest research degree.

I would like to thank the associate examiners, Professor Yuji Sakamoto (Media Creation Methodology Laboratory, Faculty of Information Science and Technology, Hokkaido University), Professor Miki Haseyama (Laboratory of Media Dynamics, Faculty of Information Science and Technology, Hokkaido University), Professor Yoshinori Dobashi (Laboratory of Information Media Environment, Faculty of Information Science and Technology, Hokkaido University), and Associate Professor Toshihiko Ito (Language Media Laboratory, Faculty of Information Science and Technology, Hokkaido University).

I would also like to express my appreciation to Assistant Professor Rafal Rzepka (Language Media Laboratory, Faculty of Information Science and Technology, Hokkaido University), for the dedicated support and guidance. Dr. Rzepka continuously provided encouragement and was always willing to discuss and give insightful advice about my research.

I would like to thank all the people I met in Sapporo City, Hokkaido, and in Yokosuka City, Kanagawa, who supported me by taking part in many experiments. Without them, I could not make progress in my research.

Special thanks to my father, Shigeru Urabe, who taught me true patience by asking me about my progress daily and nearly causing me to have a nervous breakdown. Although it was stressful, it helped me sit in front of my computer and pursue my research. Also, many thanks to my mother and sister for not asking me about my research progress ;).

Finally, my deep and sincere gratitude to my husband, Julius Michaelis, for his continuous technical support and love. He supported me fully by being an administrator for our server, “PRIDE”, and shared his coding and machine learning knowledge which made me able to finish this long and tough journey. Vielen dank (⌒3⌒)!

Research achievements

Journal Papers

1. Yuki Urabe, Rafal Rzepka and Kenji Araki: “Emoticon Recommendation System to Richen Your Online Communication”, International Journal of Multimedia Data Engineering and Management, 5(1), pp. 14-33, 2014.

2. Yuki Urabe, Rafal Rzepka and Kenji Araki: “Find Right Countenance for Your Input ー Improving Automatic Emoticon Recommendation System with Distributed Representations”, Journal of Information Processing and Management (IPM), Vol. 58, Issue. 1, 102414, 2021.

Peer-reviewed International Conference Proceedings and Lecture Notes

1. Yuki Urabe, Rafal Rzepka and Kenji Araki: “Emoticon Recommendation for Japanese Computer-Mediated Communication”, in the Proceedings of 2013 IEEE Seventh International Conference on Semantic Computing, pp. 25-31, Irvine, USA, 2013.

2. Yuki Urabe, Rafal Rzepka and Kenji Araki: “Emoticon Recommendation System for Effective Communication”, in the proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013), pp. 1460 - 1461, Niagara Falls, Canada, 2013.

3. Yuki Urabe, Rafal Rzepka and Kenji Araki: “Comparison of Emoticon Recommendation Methods to Improve Computer-Mediated Communication”, Recommendation and Search in Social Network, Lecture Notes in Social Networks, pp. 23-39, 2015.

National Conferences and Meetings

1. 卜部 有記, ジェプカ ラファウ, 荒木 健治: “顔文字の表す感情を用いた顔文字推薦システムの構築”, 言語処理学会第 19 回年次大会発表論文集, pp. 648-651, 2013.

2. 卜部 有記，ジェプカ ラファウ，荒木 健治: “ユーザの履歴選択と文の感情に基づいた顔文字推薦手法”，第 29 回ファジィシステムシンポジウム講演論文集，pp. 768-773，2013.

3. 卜部有記, ジェプカ ラファウ, 荒木 健治: “テキストメッセージの感情及びユーザの選択履歴に対応した顔文字推薦手法の提案とスマートフォンに実装した際のユーザビリティ評価”, 言語処理学会第 20 回年次大会予稿集, pp. 1119-1122, 2014.

4. 卜部 有記，ジェプカ ラファウ，荒木 健治: “ユーザの顔文字選択のプロセスからみた顔文字推薦手法の検討”, 人工知能学会全国大会論文集 第 30 回全国大会, 3H3OS17a2-3H3OS17a2, 2016.

5. 卜部 有記，ジェプカ ラファウ，荒木 健治: “複数の感情の種類と選択履歴に基づいた顔文字推薦システムにおける推薦精度低下の要因分析”, 平成 30 年度電気・情報関係学会北海道支部連合大会, pp. 186-187, 2018.

References

[1] "Copy & Paste Emoticons [in Japanese]." `https://kaomoji-copy.com/`, 2012 (accessed October 18th, 2020).

[2] F. Barbieri, F. Ronzano, and H. Saggion, "What Does This Emoji Mean? A Vector Space Skip-gram Model for Twitter Emojis," *Proceedings of the 10th International Conference on Language Resources and Evaluation*, pp. 3967–3972, 2016.

[3] F. Barbieri, J. Camacho-Collados, F. Ronzano, L. Espinosa-Anke, M. Ballesteros, V. Basile, V. Patti, and H. Saggion, "SemEval 2018 Task 2: Multilingual Emoji Prediction," *Proceedings of the 12th International Workshop on Semantic Evaluation*, pp. 24–33, 2018.

[4] Y. Urabe, R. Rzepka, and K. Araki, "Comparison of Emoticon Recommendation Methods to Improve Computer-Mediated Communication," *Recommendation and Search in Social Network, Lecture Notes in Social Networks*, pp. 23–39, 2015.

[5] Y. Urabe, R. Rzepka, and K. Araki, "Emoticon Recommendation System to Richen Your Online Communication," *2014 International Journal of Multimedia Data Engineering and Management*, vol. 5, no. 1, pp. 13–33, 2014.

[6] Y. Urabe, R. Rzepka, and K. Araki, "Find Right Countenance for Your Input — Improving Automatic Emoticon Recommendation System with Distributed Representations," *Journal of Information Processing and Management (IPM)*, vol. 58, Issue. 1, 2021.

[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 4171–4186, 2019.

[8] Y. Urabe, R. Rzepka, and K. Araki, "Study of Emoticon Selection Process for Future Emoticon Recommendation Systems [in Japanese]," *The 30th Annual Conference of the Japanese Society for Artificial Intelligence*, pp. 1–4, 2016.

[9] A. Jibril and M. Abdullah, "Relevance of Emoticons in Computer-mediated Communication Contexts: An Overview," *Asian Social Science*, vol. 9, pp. 201–207, 2013.

[10] A. J. Fridlund, "Human Facial Expression: An Evolutionary View," *Academic Press*, pp. 1–369, 1974.

[11] A. Mehrabian, "Silent Messages (1st ed)," *Wadsworth Pub. Co*, pp. 1–152, 1971.

[12] J. B. Walther and K. P. D'addario, "The Impacts of Emoticons on Message Interpretation in Computer-mediated Communication," *Social Science Computer Review*, vol. 19, no. 3, pp. 324–347, 2001.

[13] M. Ptaszynski, "Robust NLP for Real-world Data: 1. Emoticon Processing - How to Non-trivially Make a Computer Understand Trivial Expression [in Japanese]," *Information Processing Society of Japan*, vol. 53, no. 3, pp. 204–210, 2012.

[14] A. Wei, "Emoticons and the Non-verbal Communication: With Reference to Facebook (Unpublished Master' s Thesis)," *Department of Media Studies, Christ University, Bangalore UTF2013 India*, 2013.

[15] D. Derks, A. E. Bos, and J. Grumbkow, "Emoticons and Social Interaction on the Internet: The Importance of Social Context," *Computers in Human Behavior*, vol. 23, no. 1, pp. 842–849, 2007.

[16] Y. Yamamoto, T. Kumamoto, and A. Nadamoto, "Multidimensional Sentiment Calculation Method for Twitter Based on Emoticons," *International Journal of Pervasive Computing and Communications*, vol. 11, pp. 212–232, 2015.

[17] T. Harada, "The Role of Face Marks in Promoting Smooth Communication and Expressing Consideration and Politeness in Japanese [in Japanese]," *Language and Culture: the Journal of the Institute for Language and Culture*, vol. 8, pp. 205–224, 2004.

[18] S.-K. Lo, "The Nonverbal Communication Functions of Emoticons in Computer-mediated Communication," *Cyberpsychology & Behavior*, vol. 11, no. 5, pp. 595–597, 2008.

[19] J. Gajadhar and J. Green, "The Importance of Nonverbal Elements in Online Chat," *EDUCAUSE Quarterly*, vol. 28, no. 4, pp. 63–64, 2005.

[20] A. Ip, "The Impact of Emoticons on Affect Interpretation in Instant Messaging." `http://www.amysmile.com/doc/emoticon_paper.pdf`, 2002.

[21] J. Park, V. Barash, C. Fink, and M. Cha, "Emoticon Style: Interpreting Differences in Emoticons Across Cultures," *Proceedings of the 7th International Conference on Weblogs and Social Media*, pp. 466–475, 2013.

[22] M. K. Tomic, M. Martinez, and T. Vrbanec, "Emoticons," *FIP - Financije i Pravo (Journal of Finance and Law)*, vol. 1, pp. 35–42, 2013.

[23] A. Pollack, "Happy in the East (^_^) or Smiling in the West :)," *The New York Times*, vol. Section D, p. 5, 1996.

[24] B. Ashcraft, "The Latest Japanese Emoticons are Freaky, Yet Wonderful. KOTAKU." `https://kotaku.com/the-latest-japanese-emoticons-are-freaky-yet-wonderful-5941811`, 2012 (accessed June 14th, 2020).

[25] Y. Urabe, R. Rzepka, and M. Ptaszynski, "A Survery on Emoticon Recommendation Technology [in Japanese]," *Artificial Intelligence*, vol. 32, no. 3, pp. 356–363, 2017.

[26] S. Fahlman, "Smiley Lore." https://www.cs.cmu.edu/~sef/sefSmiley.htm, 2002 (accessed October 18th, 2020).

[27] Y. Wakabayashi, "Origin of Emoticons [in Japanese]." http://deafcomic.jp/Others/FaceMark.html, 2006 (accessed October 18th, 2020).

[28] "Simeji [in Japanese]." https://simeji.me/, 2013 (accessed October 18th, 2020).

[29] S. Kato, Y. Kato, and D. Scott, "Relationships Between Emotional States and Emoticons in Mobile Phone Email Communication in Japan," *International Journal on e-Learning*, vol. 8, no. 3, pp. 385–401, 2009.

[30] N. Suzuki and K. Tsuda, "Express Emoticons Choice Method for Smooth Communication of e-Business," *Knowledge-based Intelligent Information and Engineering Systems*, pp. 296–302, 2006.

[31] Y. Emura and Y. Seki, "Facemark Recommendation Based on Emotion, Communication, and Motion Type Estimation in Microblogs [in Japanese]," *Journal of Natural Language Processing*, vol. 19, no. 5, pp. 401–418, 2012.

[32] D. Derks, A. E. Bos, and J. Grumbkow, "Emoticons in Computer-mediated Communication: Social Motives and Social Context," *Cyberpsychology & Behavior*, vol. 11, no. 1, pp. 99–101, 2008.

[33] M. Inoue, M. Fujimaki, and S. Ishizaki, "System for Analyzing Emotional Expression in E-mail Text: Collection, Classification, and Analysis of Emotional Expressions [in Japanese]," *IEICE Technical Report. Thought and Language*, vol. 96, no. 608, pp. 1–8, 1997.

[34] K. Skovholt, A. Gronning, and A. Kankaanranta, "The Communicative Functions of Emoticons in Workplace E-mails: :-)," *Journal of Computer-mediated Communication*, vol. 19, pp. 780–797, 2014.

[35] A. Wolf, "Emotional Expression Online: Gender Differences in Emoticon Use," *Cyberpsychology & Behavior*, vol. 3, no. 5, pp. 827–833, 2000.

[36] Ç. Çöltekin and T. Rama, "Tübingen-Oslo at SemEval-2018 Task 2: SVMs Perform Better than RNNs in Emoji Prediction," *Proceedings of the 12th International Workshop on Semantic Evaluation*, pp. 34–38, 2018.

[37] L. Zhang, Y. Zhou, T. Erekhinskaya, and D. Moldovan, "Emoji Prediction: A Transfer Learning Approach," *Advances in Information and Communication*, pp. 864–872, 2020.

[38] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. De Jong, and U. Kaymak, "Exploiting Emoticons in Sentiment Analysis," *SAC '13: Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pp. 703–710, 2013.

[39] M. Ptaszynski, P. Dybala, R. Rzepka, and K. Araki, "Towards Fully Automatic Emoticon Analysis System (^o^)," *Proceedings of the 16th Annual Meeting of the Association for Natural Language Processing (NLP 2010)*, pp. 345–348, 2010.

[40] M. Ptaszynski, P. Dybala, R. Rzepka, K. Araki, and F. Masui, "ML-Ask: Open Source Affect Analysis Software for Textual Input in Japanese," *Journal of Open Research Software*, vol. 5, no. 1, pp. 1–16, 2017.

[41] A. Nakamura, "Dictionary of Emotive Expressions (*Kanjyo Hyogen Jiten*) [in Japanese]," *Tokyodo Publishing*, pp. 1–464, 1993.

[42] S. Kawakami, "The Database of 31 Japanese Emoticon with Their Emotions and Emphases [in Japanese]," *The Human Science Research Bulletin of Osaka Shoin Woman's University*, vol. 7, pp. 67–82, 2008.

[43] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, "The Measurement of Meaning," *Urbana, IL: University of Illinois Press*, pp. 1–342, 1957.

[44] M. Inoue and T. Kobayashi, "The Research Domain and Scale Construction of Adjective-pairs in a Semantic Differential Method in Japan," *The Japanese Journal of Educational Psychology*, vol. 33, no. 3, pp. 253–260, 1985.

[45] S. Yata, "Nihongo Web Corpus 2010." `http://www.s-yata.jp/corpus/nwc2010/`, 2010 (accessed March 1st, 2020).

[46] M. Ptaszynski, P. Dybala, R. Rzepka, K. Araki, and Y. Momouchi, "YACIS: A Five-Billion-Word Corpus of Japanese Blogs Fully Annotated with Syntactic and Affective Information," *Proceedings in the AISB/IACAP World Congress*, pp. 40–49, 2012.

[47] "Kaomoji cafe [in Japanese]." `https://kaomoji-cafe.jp/`, 2004.

[48] C. Corinna and V. Vladimir, "Support-Vector Networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[49] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed Representations of Words and Phrases and Their Compositionality," *Advances in Neural Information Processing Systems 26*, pp. 3111–3119, 2013.

[50] M. Schuster and K. Paliwal, "Bidirectional Recurrent Neural Networks," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2673–2681, 1997.

[51] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical Attention Networks for Document Classification," *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1480–1489, 2016.

[52] M. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep Contextualized Word Representations," *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 2227–2237, 2018.

[53] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature 521*, pp. 436–444, 2015.

[54] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, N. Gomez, A, L. Kaiser, and Polosukhin, "Attention is All You Need," *Advances in Neural Inforamtion Proscessing System*, vol. 30, pp. 5998–6008, 2017.

[55] S. Chen, M. Wang, and L. He, "EmojiIt at SemEval-2018 Task 2: An Effective Attention-based Recurrent Neural Network Model for Emoji Prediction with Characters Gated Words," *Proceedings of the 12th International Workshop on Semantic Evaluation*, pp. 423–427, 2018.

[56] D. Li, R. Rzepka, M. Ptaszynski, and K. Araki, "Emoji-Aware Attention-based Bi-directional GRU Network Model for Chinese Sentiment Analysis," *IJCAI Workshop on Linguistic and Cognitive Approaches to Dialog Agents Workshop LACATODA*, pp. 11–18, 2019.

[57] Q. Li, S. Li, J. Hu, S. Zhang, and J. Hu, "Tourism Review Sentiment Classification Using a Bidirectional Recurrent Neural Network with an Attention Mechanism and Topic-Enriched Word Vectors," *Sustainability*, vol. 10, no. 3313, pp. 1–16, 2018.

[58] H. Steck, "Evaluation of Recommendations: Rating-prediction and Ranking," *Proceedings of the 7th ACM Conference on Recommender Systems*, pp. 213–220, 2013.

[59] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[60] T. Kudo, K. Yamamoto, and Y. Matsumoto, "Applying Conditional Random Fields to Japanese Morphological Analysis," *Proceedings in the 2004 Conference of Empirical Methods in Natural Language Processing*, pp. 230–237, 2004.

[61] K. S. Jones, "A Statistical Interpretation of Term Specificity and its Application in Rretrieval," *Journal of Documentation*, vol. 28, no. 1, pp. 11–21, 1972.

[62] R. Řehůřek and P. Sojka, "Software Framework for Topic Modelling with Large Corpora," *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 45–50, 2010.

[63] M. Suzuki. `https://github.com/singletongue/WikiEntVec/releases/jawiki.all_vectors.300d.txt.bz2`, 2019 (accessed March 1st, 2020).

[64] Hottolink, "Word Distributed Representation Model from the Largest Japanese SNS and Web Corpus [in Japanese]." `https://github.com/hottolink/hottosns-w2v`, 2018 (accessed March 1st, 2020).

[65] D. Kingma and J. Ba, "ADAM: A Method for Stochastic Optimization," *Proceedings of the 3rd International Conference for Learning Representations*, pp. 1–15, 2015.

[66] Y. Kikuta, "BERT Pretrained Model Trained on Japanese Wikipedia Articles." `https://github.com/yoheikikuta/bert-japanese`, 2019.

[67] T. Sakaki, S. Mizuki, and N. Gunji, "BERT Pre-trained Model Trained on Large-scale Japanese Social Media Corpus." `https://github.com/hottolink/hottosns-bert`, 2019.

[68] T. Kudo and J. Richardson, "SentencePiece: A Simple and Language Dependent Subword Tokenizer and Detokenizer for Neural Text Processing," *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 66–71, 2018.

[69] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Breakthroughs in Statistics*, pp. 196–202, 1992.

[70] L. P. Morency and T. Baltrušaitis, "Multimodal Machine Learning: Integrating Languate, Vision, and Speech," *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, pp. 3–5, 2017.

[71] A. Ratner, P. Varma, B. Hancock, and C. Ré, "Weak Supervision: A New Programming Paradigm for Machine Learning." `http://ai.stanford.edu/blog/weak-supervision/`, 2019.