UNIVERSITY
of HAWAIʻI®
MĀNOA

Predicting Emoji Usage for Emoji Recommender System

Mariam Doliashvili

University of Hawaii at Manoa
mariamd@hawaii.edu

Abstract

Emojis are used frequently in social media and private conversations. They are significant means of communication that help us express emotions and describe objects visually. Previous studies have shown positive impact of emojis in human relations, memorization and user engagement with web content. Unicode version 6 includes 2923 emojis, which makes it hard to make full use of them without a recommender system. We formulate recommending emojis as a complex prediction problem based on its diverse usage as a word and as a sentiment marker.

People have individual usage of emojis, and different representation of emojis across different platforms also leads to different interpretations based on device.

Therefore, we introduce a recommender system that is able to suggest various emojis and apply personalization to increase the accuracy of the recommending process.

Exploring whether it is possible or not to extract knowledge from emoji datasets and using it to predict emoji usage, we implemented several baseline models and trained Long Short-Term Memory (LSTM) recurrent neural networks.

I. Introduction

i. Relevance

Emojis are widely used on social media sites like Twitter, Facebook and Instagram to enrich the conversation with emotions, elaborate the meaning in fewer words and so on. Their use, however, is not limited to social media: Chat groups, mail and text messages are examples of other areas where emojis are extensively used. It is an useful, fast and easy to use tool to describe one's emotions and express them in online communications. First emojis appeared on Japanese mobile phones in 1990s. the word emoji [25] comes from Japanese e (絵, "picture") + moji (文字,"character"). Their popularity increased internationally in the past two decades because they are very relevant to our everyday lives.

We like texts with emojis more and apparently using emojis also makes us more likable. According to an AMEX OPEN Forum infographic[3], emojis can make a big difference to "post" engagement rates. Posts with emojis get 33% more comments, they are shared 33% more often and they get liked 57% more often than posts without emojis. A study of the traits of highly influential social media profiles by Simo Tchokni et al. [21] showed emoticons as one of the common factors - powerful users tend to use emoticons often.

In addition, emojis have been shown to lead to better memorization of content [10]. However, there are particular cases of emoji usage where people argue that emoji usage does not always give a positive outcome.

ii. Can using emojis be negative?

It is somewhat ambiguous, whether one should use an emoji at work related emails or not. Jina Yoo [28] tested how people perceive smiley faces in a work email as compared to a social email. Researchers sent two types of email messages to a group: a flirtatious message, and another one about extending a job interview request. Emoticons were added to some texts of each type. "The usage of emoticons induced stronger relational outcomes in a task-oriented context than in a socio-emotional context. Participants in a task-oriented condition actually liked the sender more if the sender used an emoticon rather than if the sender used no emoticons" and the sender's credibility wasn't affected by the emoticons even when they used up to 4 emoticons. As the possible explanation of the result is given the following: "emoticons are overused already in socio-emotional contexts, and no special value is assigned to using emoticons in email in the same context. However, when the emoticons are used in a task-oriented context, they might function as a positive expectancy violation, which could bring positive relational outcomes." Contrary to this, Glikson at al[14] published a paper "The Dark Side of a Smiley, Effects of Smiling Emoticons on Virtual First Impressions" where it is stated that "contrary to actual smiles, smileys do not increase perceptions of warmth and actually decrease perceptions of competence" and "Perceptions of low competence, in turn, undermined information sharing." (The authors also mentioned that if all the team members were younger, likelihood of using emoticons in the team's conversation was higher). Taking into account that both of the studies are evaluated in only particular scenarios (1 - extending a job interview request, 2 - First impressions over internet) we cannot draw a general conclusion.

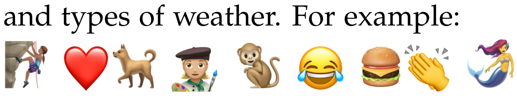
One thing to note is that most of these papers study only emoticons not visual emojis. Since emojis were available for large public from 2010.

However, Wang et al. [26] showed that emoticons reduced the negativity effect in the business-related email messages - the same message sounded less negative when paired with a positive (smiley) emoticon.

In addition Kalyanaraman et al. [17] conducted a study that had participants chat online with "health experts" and "film experts" who either used or avoided emoticons, the participants rated the experts in both topics friendlier and more competent when they communicated with emoticons. This study also noted that emoticons might help you remember what you've read more easily - "It appears that the presence of emoticons affects cognition as well, because participants" scores on memory for chat content were significantly higher in the "emoticons present" condition than in the "emoticons absent" condition."

iii. History

Emojis are often confused with emoticons. An emoticon is a representation of a facial expression using punctuation marks, numbers and letters, usually written to express a person's feelings or mood, e.g. :) :P (-_-) ^_^ xD. While emojis are used like emoticons they are small digital images or icons that exist in various genres, including facial expressions, common objects, food, activities, animals, places and types of weather. For example:



"The development of emoji was predated by text-based emoticons, as well as graphical representations, inside and outside of Japan. Scott Fahlman, a computer scientist at Carnegie Mellon University was credited with popularizing early text-based emoticons in 1982, when he proposed using the : -) and : - (sequence of characters to help readers of a school message board distinguish between serious posts and jokes. The first emoji was created in 1999 in Japan by Shigetaka Kurita." [4] From 2010 onwards, hundreds of emoji character sets have been incorporated into Unicode, a standard system for indexing characters, which has allowed them to be used outside Japan and to be standardized across different operating systems. After 2010 each update of the unicode standards introduced new sets of emojis. "Corporate demand for emoji standardisation has placed pressures on the Unicode Consortium, with some members complaining that it had overtaken the group's traditional focus on standardising characters used for minority languages and transcribing historical records." [4]

II. Difficulties of building an emoji recommender system

The fact that emojis are related to emotions and are becoming means of communication makes emoji prediction an interesting problem for Natural Language Processing (NLP). If we assume that an emoji is a label of the text corresponding to an emotion then we would face the sentiment analysis problem. However the classical sentiment analysis problem only analyses whether a text is positive or negative - sentiment polarities of sentences and documents. Advanced models only have several additional emotions like happiness and anger. On the other hand, emoji classification has larger amount of candidates. As for now - Nov 2017 - there are 2623 unicode emojis available [16]. They are much more detailed and complicated to predict, because one emoji corresponds to many emotions based on a use case and the same emotion can be expressed with various emojis. So sentiment analysis using emojis as emotional markers would make a tedious task to solve. Nowadays emojis are not expressing only emotions, they also describe professions, activities, flags, food and even fairies and mermaids.They have gender and racial diversity. Such diversity and amount of emojis makes it more necessary to have a recommender system.

Furthermore recommending emojis in a chat application also requires the understanding of a conversation, since an emoji can be used as:

- An answer to the previous text if it was a question,
- A reply for the previous text,
- A next word in the sentence.

Taking this into account if we were to fully cover the task of emoji prediction we would also need to address question-answering, smart reply and a next word prediction problems. We only cover the next word prediction because of the following reasons: It is hard to determine a start and an endpoint of a conversation and we also need to take into account that several people that can contribute to the conversation might have completely different emoji usage and conversation style. For instance, a conversation is just two texts: #1 - from Bob to Alice and #2 - from Alice to Bob. #1 is a lengthy text, includes emojis and they are mostly negative. Alice never uses negative emojis.

Assuming that people would use the same emojis in the same situations is a mistake that we discuss later.

III. Description of the task

The problem formulation that we are trying to solve is the following: Predict the top K emojis that Alice would use after writing a short message that may or may not already include emojis.

The Facebook sentiment analysis paper by Tial et al. [23] shows that a sentence with an emoji does not necessarily equal to the same sentence without emojis. Therefore we need to take into account the words in a sentence and the emojis in the text.

To begin with, studying the Twitter dataset shows three main cases of emoji usage with a text (two of which we already discussed):

- Expressing an emotion about the text it accompanies.
  I.e. "Last couple months have been crazy! 😭😭"
- As a word in the sentence.
  I.e. "I ❤️ you", "Damn, I would love this. Or suicide squad, working towards that. Patty was 👌💯"
- Emoji combinations to express...
  - ... a phrase - "☕🕓❓"- what time should we get coffee?. "🚖🛂✈❓"- Do you need a ride to the airport?. "Had an accounting midterm today that I wasn't expecting until next week. Lots of unfamiliar terms 😢🔫. Here 😢🔫 stands for 'kill me now'.

– ... the strength of emotion - "😥😢😭"- very sad. "👏 👏 👏👏👏"- bravo. 😂😂😂- very funny. etc.
– ... an emotion/entity that has no separate emoji in the unicode yet. ?With me now: Every damn corner 🌮🚐 #TacoTuesday?. Here 🌮🚐 stands for a taco food truck. "🐓🐔🍳🐤🐣"- life cycle. "Anti-gravity 🌆🏃" - 🌆🏃 stands for 'city parkour'.

We are able to study the given sentence for the above aspects, however there are many other factors besides the actual text that affect the usage of emoji. The relations between emojis and emotions are quite ambiguous themselves, given the fact that there are many emojis [12] to express the same emotion and the same emoji can be used to express different emotions [18]. Furthermore the same emoji can have very different meanings for different people and it can lead to misunderstandings [24]. Especially, this becomes a more obvious problem if we look at the different representations of the same unicode emoji across various platforms. Miller et al. [18] also shows how different the same unicode emojis can be perceived even for the same person. Therefore people can change their emoji usage based on the device they use at the moment of typing.

Various studies over the years show that emoji usage can be different based on gender (females are using emojis with tears more frequently [2]) and culture (Different countries have different favorite emojis [1]). Emoji usage can also depend on the location of the person texting. For instance, a person in Hawaii is more likely to use palm tree and pineapple than a person in Siberia. However if Hawaiian residents decide to visit Mauna Kea (is a dormant volcano, its peak is the highest point in the state of Hawaii) in winter they are also likely to use snow/winter emojis. Therefore, emoji usage can also depend on the weather in the area. The location and nationality of the person also gives a good prediction of the flags they are going to use. Based on the dataset we can also get an intuition that people in the same friend circle tend to use similar emojis, like they tend to use similar vocabulary. Another interesting observation is that people tend to use more emojis with short texts than with larger texts (probably because they are giving so much context in the text with words that emojis are not necessary). Finally, age also can be an interesting feature to observe, not only because there are differently aged emojis, but also people in different generations view emojis differently.

All the above differences make it almost impossible for the knowledge drawn from a general dataset to make a prediction for an individual with no further background knowledge. Because of this we have decided to personalize the training for emoji recommendation. However, we need to take into account that given the above features (age, gender, location, culture) about users it can be possible to profile them and use similar profiles to train/predict together. Since the features are not accessible for us at the time being, we would like to leave it as a future work. (It would solve the following case - What happens if people do not already use emojis?)

i. Contribution

To the best of our knowledge we are first to personalize emoji predictions and there is no existent model before this that does not limit the existing set of unicode emojis to a smaller subset. In addition we divide the emoji prediction problem as 'sentiment analysis', 'next word prediction' and 'word - emotion mapping' tasks and introduce a recommender system that combines the three models with a heuristic to give a recommendation of the top k most relevant emojis.

IV. Related Work

The first thing that worth noting is that all the related work that we found with emoji predictions uses emojis as a sentiment marker. A recent study from Barbieri et al. [5] addresses the prediction problem using a multi class LSTM classifier to predict one emoji per tweet. They have also mentioned that their system predictions are more accurate than human evaluators, since it is not easy for humans to predict emoji from the text without background knowledge about the text and the person. This, combined with the fact that individuals have a different understanding of the same emojis [18] brings up an interesting perspective that predictions are dependent on individual people and their way of writing and expressing emotions. In addition, the same study [18] shows that because of the different representations of emojis across different platforms people assign various emotions to the same unicode emoji. Therefore, all the above indicate that it is a good idea to personalize emoji recommendations taking into account "individual, device" as a separate entity. Since sentiment analysis is an interesting subject to research for NLP, there has been a vast amount of research about the topic. However, there are not many resources about emoji usage and predictions. The above mentioned Barbieri et al. [5] addresses the prediction problem to predict one emoji per tweet but they only attempt to predict X different emojis. As for now, to the best of our knowledge there is no article that analyses data using all 2623 emojis. All of the papers we come across about emoji predictions [5] [27] limit the amount of emojis below 65 of the most used emojis. Most of the models with good performance (accuracy > 50%) [5] [27] only classify below 10 emojis. Since some of the platforms, by default, sort the emoji list putting the most used ones on the top, recommending only one of top k emojis is not helping users much. At the same time, recommending emojis that are not frequently used because they are hard to find and would be relevant to the text is likely to increase user satisfaction and emoji popularity.

V. Dataset and Methods

For the study, four datasets have been collected from Twitter, including 600,000 tweets (Dataset #1) and 50,000 tweets (Dataset #2) for the following 74 emojis:

'😁', '😂', '😃', '😄', '😅', '😆', '😉',
'😓', '😔', '😖', '😘', '😚', '😜', '😝',
'😥', '😨', '😩', '😪', '😫', '😭', '😰',
'😹', '😺', '😻', '😼', '😽', '😾', '😿',
'🙊', '🙋', '🙌', '🙍', '🙎', '🙏', '🤔',
'💁', '👌', '👏', '💕', '❤️', '💗', '🔥',
'😊', '😋', '😌', '😍', '😏', '😒', '🇺🇸',
'😞', '😠', '😡', '😢', '😣', '😤', '♥️',
'😱', '😲', '😳', '😵', '😷', '😸', '👩',
'🙀', '🙅', '🙆', '🙇', '🙈', '🙉', '🤣',
'🤗', '😎', '😋', '😆'

The third dataset (Dataset #3) has been collected for a limited amount of emojis and amounts to 50,000 tweets:

'😂', '😉', '😊', '😍','😘',
'😜', '😞', '😡', '😭', '😱'

For studying personalized emoji recommendations we used complete datasets of 5 users that were willing to participate in the study.

i. Methods

For experimenting purposes we have designed models for the following formulations of tasks:

- The initial one is to predict an emoji that is used in the tweet, in any possible position. It is based on the general dataset - the tweets are gathered from multiple users in a 5-10 hour time frame.
- Taking into account how complex and diverse the above problem is, we have also tried to simplify the prediction by making it a binary classification problem: For a chosen emoji we tried to predict whether it will be used in the tweet or not.
- Finally, predict an emoji that is used after the given text, based on one user?s data.
- The fact that one emotion can be described with various emojis lead us to think that for better results we should recommend a set of emojis that correspond to the same emotion. With this modification the binary classification problem changes to predict whether a tweet contains an emoji from the given set and for the general prediction we can recommend top k emojis based on their probabilities.

ii. Dataset

For not personalized predictions the data was collected using the Twitter Streaming API [11] in the time period of April 29 - May 1, 2017. The language used in the dataset is English. The only preprocessing that has modified the dataset was removing the unnecessary information: URLs, #-signs, malformed words containing numbers, etc. However, from the fact that the data contains a significant amount of spelling errors, it is worth mentioning that using spelling correction might increase the quality of the dataset and lead to a better solution in the future.

The above datasets are labeled in the two different ways:

- Each tweet labeled with the emoji used in it.

- Each tweet labeled with 1/0 based on whether or not a specific emoji/or an emoji from a specific set of emojis is used in the tweet.

The second case also needs balancing the dataset after the labeling, since it can result in a disproportionate amount of data for a binary classification problem. One more thing to mention is that there are many cases when a tweet contains several emojis. In this case, we preferred to label the tweet with the first occurring emoji. It is optimal for this situation, because the emoji is by default one of the labels of the tweet and the choice of the first emoji is random; in addition labeling the same tweet with various labels would create a problem that we would have to address later on the learning phase.

For personalized prediction the data for training, testing and evaluation is from the same Twitter user. It includes 47000 tweets.

For labeling the dataset we use all available unicode emojis. We created a mapping of each emoji to its description and keywords that it associates with. Tweets are split into sentences that are followed by an emoji and then labeled with the emoji that was following the text.

E.g. I like that <3 thought I would not participate :/
Would produce the following sentences and labels:
Sentence 1: I like that
Label 1: <3
Sentence 2: I like that <3 thought I would not participate
Label 2: :/

In addition, we took the combinations of emojis that were used together to express a non-existent emoji or a phrase (in short we call them combojis) into account and added them into the labels' list. For each label we calculate and update the frequency of how many times it is used so far. When combojis frequency reaches to certain limit we generate a mapping and keywords for it, based on the texts that it was used with in the past. The frequency is also used as a feature for training.

iii. Word representations

For the word representations, we used one hot encoding and word embeddings.

- **One Hot Encoding.** For one hot encoding [8], we calculate the frequency of words in all tweets. Then, we take the k most frequently used words and create a binary vector for each tweet. Each binary vector has 5,000 entries, where an entry corresponds to one of the k most frequent words. Entries are filled with an 1, if the corresponding word is present in the tweet and 0 if it is not.
- **Word Embeddings.** For word embeddings [6], we, again, calculate the frequency of words in all tweets. For each tweet, we create a vector with as many entries as there are words in a tweet. Each entry is filled by using a word's index in the k most frequently used words as the value. If a word does not occur in the k most frequent words, we fill the vector entry with zero. In the end, we train the word embeddings with the neural network. Similar words should be placed close together in the vector space after the learning. Finally, the resulting word representations are split up into training, testing and evaluation sets.

iv. Personalized learning

The model for personalized learning combines in itself three solutions for subtasks and a scoring function. As described in the introduction, emojis and combojis can be expressing an emotion and a next word in the sentence. Therefore we combined approaches of next word suggestion algorithm and sentiment analysis.

v. Next word suggestion algorithm

We need to construct an algorithm that fulfills the following steps: Build a language model using twitter text and then use this language model to predict the next word as a user types.

We need to calculate the frequency of words and n-grams and use a sliding window to do so.

If we assume the training data shows the frequency of "university" is 198, "university student" is 12 and "university professor" is 10. We calculate the maximum likelihood estimate (MLE) as:

- The probability of "university student":
  $P_{mle}(entry|data) = 12/198 = 0.06 = 6\%$
- The probability of "university professor" is:
  $P_{mle}(streams|data) = 10/198 = 0.05 = 5\%$

If the user types, "university", the model predicts that "student" is the most likely next word.

The n-gram model description steps are:

- Generate 2-grams and 3-grams.
- Select n-grams that account for 60% of word instances. This reduces the size of the models.
- Calculate the maximum likelihood estimate (MLE) for words, for each model.

As for the prediction: We use the ngram models on tokenized and preprocessed user input. We implement a Stupid backoff [7] starting on the 3-gram model backing off to the 2-gram model and returning 3 words with the largest MLE

vi. Mapping from a word to emoji

We created the mapping using 2623 available unicode emojis from unicode.org and their descriptions as names. We generate keywords based on the frequency of the words that they are used with in a sentence. The mapping file is later updated to store informations for combojis. The file will be used to map the next predicted word to associated emojis.

vii. Evaluating emotion

Both in general and personalized tasks we train a Long Short Term Memory (LSTM) network to predict which emojis are used with the text. LSTM is a special kind of Recurrent Neural Network (RNN) that is able to learn long term dependencies. This kind of RNNs are good at remembering information for long periods of time.

LSTMs were introduced by Hochreiter and Schmidhuber [15]. However over the course of years they were popularized and developed by various contributors, since they perform well on a large variety of problems. Over the course of years, LSTMs proved to perform well for NLP tasks including sentiment analysis.[cite]

viii. Scoring

Scoring of emojis in the final stage is based on the prediction probability produced by a trained model for labeling, label frequency in the existing dataset that the model is trained on and an additional feature for measuring confidence.

Out of the above two models we get two sets of predictions for each emoji. The next word prediction task with mapping assigns probabilities to all the possible labels(emojis) - whether they have been used before or not. The sentiment prediction part only uses already used emojis as labels. Therefore it generates probabilities only for a limited set of emojis. We have a confidence score to adjust the two probabilities based on the relative frequency of emojis before.

For example: If an emoji has a prediction probability 0.9 and its relative frequency is in the top 10% we give it a high confidence score. Finally, we calculate the weighted sum of the prediction probability and confidence and recommend the top k emojis for a given text.

ix. Implementation

While achieving the results of this paper required a substantial amount of own scripting, three Python libraries were essential to the analysis.

- **NLTK [19].** The Natural Language ToolKit (NLTK) is a Python library specialized in natural language processing. We made use of its word tokenizing capabilities and used its naive Bayes classifier to create the baseline method.
- **SciKit [13]** Scikit is a machine learning library for python that offers tools for data mining and analysis. We have used its implementations of the algorithms - Logistic Regression and Stochastic Gradient Descent.
- **Keras on TensorFlow [20] [22].** Keras is a high-level neural networks API, written in Python and capable of running on top of either TensorFlow or Theano. While Tensorflow itself is a open-source library for numerical computations, developed by Google Brain Team for the purposes of conducting machine learning and deep neural networks research. We have used it to train word embedding vectors, build an LSTM classifier and optimize it for accuracy.

The NLTK and SciKit classifiers are used with the one hot encoding of the tweets. Each tweet is labeled with the emoji, that is contained in the tweet, as described above.

The LSTM neural network is used with word embeddings vectors. The dataset has been labeled in the following ways:

- For a specific emoji:
  1. The label for each tweet equals to 1 if the tweet contains the emoji.
  2. Otherwise the label equals to 0.
- For a specific emotion:
  1. Create a set of emojis that correspond to the emotion.
  2. The label for each tweet equals to 1 if the tweet contains an emoji from the given set.
  3. Otherwise the label equals to 0.

The emojis used for a single emoji classifier are:

'😂', '😉', '😊', '😍','😘',
'😜', '😞', '😡', '😭', '😱'

The emoji sets used related to the same emotion are:

1. '😁', '😂', '😃', '😄', '😆',
2. '😂', '🤣', '😃', '😄', '😆',
3. '😁', '😂', '😉', '😊', '😹',
4. '😉', '😊', '😹', '😺', '😆',
5. '😍', '😘', '😻', '💕', '❤️',

The example of combojis generated after training on an individual data are:

👉😎👉	☕🕓❓	😥😢😭
🌮🚐	🔥💜	😢🔫
👀😜	🐓🐔🐤🐥🐣	🎈🤡

VI. Evaluation methods

For the general task we use the prediction accuracy as a metric.

As for the individual prediction we have two evaluation metrics: Precision at the top k candidates and the mean reciprocal rank that is used to evaluate the ranking of the top predictions.

In the future, it is necessary to use human evaluation in the process - If people did not use emojis in a text that does not mean that they would not use them if a recommender system was available.

VII. Analysis

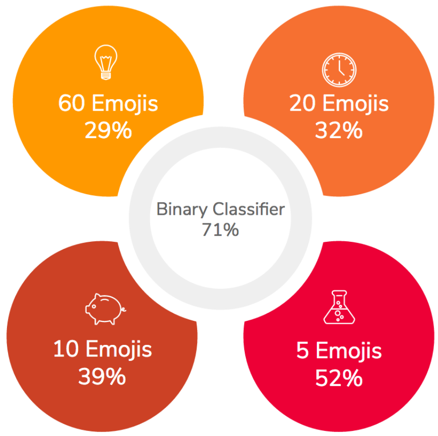
Since the computations used in the implementation are quite time costly, the classifiers for multi-variable problem were trained on the smaller dataset (Dataset #2). For the results please see the Table 1.

Table 1: *Accuracy of the naive approach*

Dataset #2			
Algorithm	5emojis	10emojis	20+emojis
Naive Bayes	48%	39%	32%
Logistic Regression	52%	38%	32%
Stochastic Gradient Descent	13%	9%	7%

Unfortunately Stochastic Gradient Descent was not a fit for the problem, as for the Naive Bayes and Logistic Regression they have shown improvements when hyper-parameters were reset. The LSTM neural network was implemented solely for the binary classification problems, using word embeddings. It has been directed to optimize the accuracy over the training phase. This resulted in an average of 71% accuracy for a single emoji classifier, and an average of 70% accuracy for the classifier of a set containing 4 emojis related to the same emotion. The emoji sets related to the same emotion were chosen naively and for the future improvements it is necessary to explore the ways to create the emojis sets that are the most related to each other and interchangeable in the everyday usage.

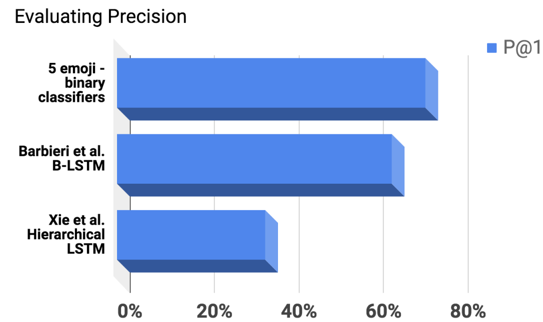
The following figure summarizes the overall evaluation for the average values of accuracy:



The binary classifier result demonstrated in the figure is trained, tested and evaluated on 6:2:2 proportions of 365 000 tweets. The original dataset used is Dataset #1, which after preprocessing and balancing is reduced to 365 000 tweets. This is quite good in our case, since we can use the binary classifier to determine for each emoji whether it should be recommended for a tweet or not. Given the result we can assume that, if recommendation for an emoji is given with 71% of accuracy, a recommender system that uses suggestion of 3 top results would significantly increase overall accuracy of the suggestion.

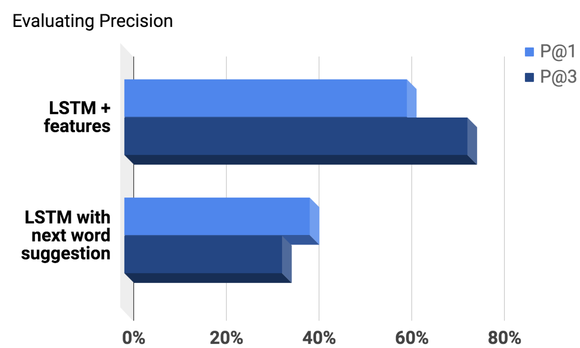
Experiments [5] showed that human evaluators on average achieve 80% accuracy on twitter dataset. Therefore accuracy of the recommendation for each emoji equals to 0.71 means that the system is quite close to how a human would perform.

To compare the results with existent paper, we created a combination of 5 binary classifiers that predicts the one emoji that has the highest probability out of 5. It seems to perform better, however we need to take into account that the datasets are not the same. Especially Xie et al. [27] uses Weibo data in Chinese language.



Evaluating personalized recommender:

Accuracy for the next word suggestion algorithm using stupid backoff is 13.5%. After mapping the words with emojis and combining it with LSTM predicted emojis recommender system had the accuracy of 34% on average. While LSTM only is able to achieve 74%.



In fact, it is expected that incorporating Next Word Prediction(NWP) algorithm decreases the performance measure in this case. Since NWP together with mapping does not limit any emojis from the full unicode emoji list [16] it leads to reach recommendations by including wide range of emojis that has not been used before. The users in our personalized dataset only use up to 125 emojis out of 2623 available. In order to better evaluate the accuracy of our recommender, in the future we need to recruit active Twitter users to be able to train on their dataset and keep track of how many recommendations they find useful. This would give us an opportunity to have a precise evaluation of the accuracy and expectations of its usage in real-world setting.

VIII. Conclusion

Emojis are used in everyday life by millions of people, which makes them a widely used tool for expressing emotions. Therefore there is a vast amount of data in the web to experiment with Machine Learning algorithms for an emoji related problem. It is an interesting NLP task to analyze and explore the ways they relate to emotions. In addition, the existence of emojis gives us the opportunity to research the emotions of the online population towards particular events, i.e. analyzing sentiment towards 2016 U.S. Presidential Candidates [9], etc. Which makes us develop systems that can make it easier for users to include emojis in their texts [29]. This paper shows that it is possible to create a recommender system for emojis with reasonable accuracy.

i. Multi-variable vs. Binary classifier

The task of creating a multi-variable classifier is easy to formulate. It is also easy to create the necessary dataset. However, it required exploring of emoji usage first: Which emojis correspond to the same emotion? How many emojis to select for a single study? Which emojis are used together? etc. The complexity of the problem makes it hard to achieve good results. In contrast, it is fairly possible to predict usage of one emoji. Solving the problem for individual cases lets us formulate the solution of the initial problem. Recommending several emojis from the same classifier also increases the accuracy of a recommender system.

Additionally there are also issues to address regarding the binary classifier. It introduces bias and needs balancing of the dataset.

Another way to resolve all the differences in emoji usage for a multi classifier that we discussed is personalizing the dataset for a particular user. This helps us to develop a system that would be highly useful for active users, however can be problematic for the users that do not produce enough data for the model to learn. Finally, for a future work, we also mentioned that generalizing task by not restricting the dataset but creating user profiles can have a potential to achieve a good result. Since it would be possible to create a personalization element for ambiguous emojis and benefit from the knowledge acquired from a larger dataset.

ii. Future Work

As future work, we can do improvements for various steps of the development:

- **Dataset**:
  - Training the model on a bigger dataset from individual users to make it possible to profile based on features like location and ethnicity.
  - We need to evaluate accuracy of the above methods compared to human operators.
- **Clustering emojis**: Exploring ways to create improved mapping (emoji:words) for the classifiers:

– Set of emotions that have corresponding emojis without intersection.
– Set of the emojis that correspond to the same emotion.

The latter would allow us to improve the error function, since recommending an emoji that is related to the correct label should be weighted as a smaller error.

Finally, we plan to test the system performance in a real world setting.

IX. Aknowledgments

Special thanks to Dr. Lipyeow Lim and Dr. David Chin for useful advices. Ed White and Gene Park for collecting and donating personal data.

[12] [19] [13] [20] [22] [11]

References

[1] Swiftkey emoji report. `https://www.scribd.com/doc/262594751/SwiftKey-Emoji-Report`, 2015. [Online; accessed 2017-12-4].

[2] Emoji gender study. `https://www.brandwatch.com/blog/react-gender-emoji-data/`, 2017. [Online; accessed 2017-12-4].

[3] Emojis effect on facebook engagement. `https://blog.bufferapp.com/\7-facebook-stats-you-should-\know-for-a-more-engaging-page`, 2017. [Online; accessed 2017-12-4].

[4] History of emojis. `https://en.wikipedia.org/wiki/Emoji#History`, 2017. [Online; accessed 2017-12-4].

[5] Francesco Barbieri, Miguel Ballesteros, and Horacio Saggion. Are emojis predictable? *CoRR*, abs/1702.07285, 2017.

[6] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Janvin. A neural probabilistic language model. *J. Mach. Learn. Res.*, 3:1137–1155, March 2003.

[7] Thorsten Brants, Ashok Popat, Peng Xu, Franz Josef Och, and Jeffrey Dean. Large language models in machine translation., 01 2007.

[8] Maico Cassel and Fernanda Lima Kastensmidt. Evaluating one-hot encoding finite state machines for seu reliability in sram-based fpgas. In *Proceedings of the 12th IEEE International Symposium on On-Line Testing*, IOLTS '06, pages 139–144, Washington, DC, USA, 2006. IEEE Computer Society.

[9] D. Chin, A. Zappone, and J. Zhaon. Analyzing twitter sentiment of the 2016 presidential candidates, 2015.

[10] M. Danesi. *The Semiotics of Emoji: The Rise of Visual Language in the Age of the Internet*. Bloomsbury Advances in Semiotics. Bloomsbury Publishing, 2016.

[11] Twitter Developers. Twitter streaming api. `https://dev.twitter.com/streaming/overview`, 2017. [Online; accessed 2017-12-4].

[12] Emojipedia. Emojis available for apple products. `https://dev.twitter.com/streaming/overview`, 2017.

[13] French Institute for Research in Computer Science and Automation (INRIA). Scikit. `http://scikit-learn.`

org/, 2017. [Online; accessed 2017-12-4].

[14] Ella Glikson, Arik Cheshin, and Gerben A. van Kleef. The dark side of a smiley: Effects of smiling emoticons on virtual first impressions. *Social Psychological and Personality Science*, 0(0):1948550617720269, 0.

[15] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, November 1997.

[16] Unicode Inc. Unicode version 5 emojis full list. https://unicode.org/emoji/charts-5.0/full-emoji-list.html, 2017. [Online; accessed 2017-12-4].

[17] S. Kalyanaraman and J. D. Ivory. The face of online information processing: Effects of emoticons on impression formation, affect, and cognition in chat transcripts, 2013-12-17.

[18] Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. *"blissfully happy" or "ready to fight": Varying interpretations of emoji*, pages 259–268. AAAI press, 2016.

[19] Team NLTK. Natural language toolkit. http://www.nltk.org/, 2017. [Online; accessed 2017-12-4].

[20] François Chollet & others. Keras: The python deep learning library. https://keras.io/, 2017. [Online; accessed 2017-12-4].

[21] Simo Tchokni, Diarmuid O Seaghdha, and Daniele Quercia. Emoticons and phrases: Status symbols in social media, 2014.

[22] Google Brain Team. Tensorflow. https://www.tensorflow.org/, 2017. [Online; accessed 2017-12-4].

[23] Ye Tian, Thiago Galery, Giulio Dulcinati, Emilia Molimpakis, and Chao Sun. Facebook sentiment: Reactions and emojis, 01 2017.

[24] Garreth W. Tigwell and David R. Flatla. Oh that's what you meant!: Reducing emoji misunderstanding. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, MobileHCI '16, pages 859–866, New York, NY, USA, 2016. ACM.

[25] Cambridge University. Emoji meaning in the cambridge english dictionary. http://dictionary.cambridge.org/dictionary/english/emoji, 2017. [Online; accessed 2017-12-4].

[26] Weiquan Wang, Yi Zhao, Lingyun Qiu, and Yan Zhu. Effects of emoticons on the acceptance of negative feedback in computer-mediated communication. 15:454–483, 08 2014.

[27] Ruobing Xie, Zhiyuan Liu, Rui Yan, and Maosong Sun. Neural emoji recommendation in dialogue systems. *CoRR*, abs/1612.04609, 2016.

[28] Jina H. Yoo. To smile or not to smile :): Defining the effects of emoticons on relational outcomes, 2007-05-24.

[29] Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu. Emotional chatting machine: Emotional conversation generation with

internal and external memory. *CoRR*, abs/1704.01074, 2017.