

IIT DHANBAD CODECHAMPS at SemEval-2022 Task 5: MAMI - Multimedia Automatic Misogyny Identification

Shubham Kumar Barnwal Department of Computer Science and Engineering , Indian Institute of Technology (Indian School of Mines), Dhanbad, India shubham.developer02@gmail.com	Ritesh Kumar Department of Computer Science and Engineering, National Institute of Technology Jamshedpur, India ritesh.cse@nitjsr.ac.in	Rajendra Pamula Department of Computer Science and Engineering , Indian Institute of Technology (Indian School of Mines), Dhanbad, India rajendrapamula@gmail.com
---	---	--

Abstract

With the growth of the internet, the use of social media based on images has drastically increased like Twitter, Instagram, etc. In these social media, women have a very high contribution as of 75% women use social media multiple times compared to men which is only 65% of men uses social media multiple times a day. However, with this much contribution, it also increases systematic inequality and discrimination offline is replicated in online spaces in the form of MEMEs. A meme is essentially an image characterized by pictorial content with an overlaying text a posteriori introduced by humans, with the main goal of being funny and/or ironic. Although most of them are created with the intent of making funny jokes, in a short time people started to use them as a form of hate and prejudice against women, landing to sexist and aggressive messages in online environments that subsequently amplify the sexual stereotyping and gender inequality of the offline world. This leads to the need for automatic detection of Misogyny MEMEs. Specifically, I described the model submitted for the shared task on Multimedia Automatic Misogyny Identification (MAMI (Fersini et al., 2022)) and my team name is IIT DHANBAD CODECHAMPS.

1 Introduction

With the growth of the internet, social media becomes a crucial part of everyone's life. As every coin has two side positive and negative, social media also comes with a number of problems. The challenge of identifying misogyny (Srivastava et al., 2017) in different social media specially in forms of meme which contains both image and text is very complicated. Misogyny meme highly affected the life of women's as its spread hate and prejudice behaviour against women's. Social media like twitter, Instagram, etc have handled by their own ways. However, detecting such memes is highly challenging. Due to this challenge, it attracts the

researcher's attention. According to one social media Instagram, more than 1 million users shared memes daily. So, with this huge amount of data in social medias and internet it is impossible to detect every misogyny meme by man power. So, we need machine learning, deep learning and artificial intelligence techniques to detect automatically misogyny memes in social media. In this paper, we have explored various Machine Learning (ML) and Deep Learning (DL) algorithms for misogyny identification in shared task MAMI (Fersini et al., 2022) challenge and my team's name is IIT DHANBAD CODECHAMPS. As per requirement of MAMI, I have submitted 4 runs for Subtask-A. My best run in Subtask-A has achieved Macro-F1 score of 0.656.

2 Related Works

Many works related to automatic detection of misogyny, hate, sexism on social media and web have been proposed.

Abir Rahali (Rahali et al., 2021) proposed a approach for automatic misogyny detection in social media using attention based bidirectional LSTM. Endang Wahyu Pamungkas (Pamungkas et al., 2020) proposed a method for Automatic Identification of Misogyny in English and Italian Tweets at EVALITA 2018 with a Multilingual Hate Lexicon. Mario Anzovino, Elisabetta Fersini (Anzovino et al., 2018) proposed a method for Automatic Identification and Classification of Misogynistic Language on Twitter. The main contribution of this paper is two-fold: (1) a corpus of misogynous tweets, labelled from different perspective and (2) an exploratory investigation on NLP features and ML models for detecting and classifying misogynistic language.

Rachael Fulper (Fulper et al., 2014) proposed a relation between misogynistic language in twitter and sexual Violence. In their paper they consider all 50 states in Washington DC.

Lakes Goenaga, Aitziber (Goenaga et al., 2018) Atutxa proposed a Automatic misogyny identification using neural networks. In this paper they focus on recurrent neural network (RNN) approach using a Bidirectional Long Short Term Memory (Bi-LSTM).

3 Task and Dataset Description

Here we have described the dataset and task provided by Multimedia Automatic Misogyny Identification (MAMI (Fersini et al., 2022)) challenge.

Multimedia Automatic Misogyny Identification (MAMI) task is divided into two sub task. Sub-task A: a basic task about misogynous meme identification, where a meme should be categorized either as misogynous or not misogynous (shown in Table 1).

Sub-task B: an advanced task, where the type of misogyny should be recognized among potential overlapping categories such as stereotype, shaming, objectification and violence. *e.g.*

1026.jpg 10101 POV: You're my wife made with mematic

POV: You're my wife



Figure 1: 1026.jpg

Here, 1026.jpg represent meme file name . Next column contains 5 numbers of zeros and ones . First numbers represent wheather meme is misogynous . Second numbers represent wheather meme is shaming . Third numbers represent wheather meme is stereotype .Fourth numbers represent wheather meme is objectification .Fifth numbers represent wheather meme is violence . Next col-

umn represent Text Transcription of the meme.

4 Methodology

4.1 Text Preprocessing

First, we removed all the punctuations, numbers, links and stop words. We have used lemmatization for grouping together the different forms of a word into a single word. NLTK wordnet (Loper and Bird, 2002) is used for lemmatization.

4.2 Feature Extraction

TfidfVectorizer (Kumar and Subba, 2020) is used for converting the text into numerical features. Pipeline ¹ is used for doing TfidfVectorizer and classification in pipelined manner. Tokenizer by keras library is used for LSTM and Bert. For Logistic regression and SVM we have used TfidfVectorizer from scikit-learn library.

Models Proposed

For Subtask-A, we have submitted 4 runs based on four different algorithms, namely- Logistic Regression (Sammut and Webb, 2010), SVM (Noble, 2006), LSTM (Hochreiter and Schmidhuber, 1997), Bert (Devlin et al., 2018) with different parameters like batch size, epochs, number of perceptron etc. We have used the scikit-learn library for logistic regression based models and SVM (support vector machines) models. Keras is used for LSTM and BERT. We scored maximum F1 score 0.656 using BERT. We have used the following value of parameters :-

- 1.For TfidfVectorizer, we have used min_df=20, max_features=2000 and max_df=0.6 .
2. For LSTM and BERT, we have used batch size = 2, epochs = 3 and number of layers = 2 .

5 Result and Discussions

The results of Subtask-A are represented in terms of Macro-F1 (shown in Table 2). The best score as Macro-F1 for Subtask-A we get is 0.656. Table 2 shows the score of our submissions based on different algorithms on MAMI challenge official ranking.

For Subtask-A BERT performs better than all other models with the parameters batch size = 2 , epochs = 3 , number of hidden layers = 2 and number of perceptron's is 128 in first layer and 64 in second layer.

¹shorturl.at/eqrFW

file_name	misogynous	Text transcription
1.jpg	0	Milk Milk.zip
10000.jpg	0	MAN SEEKING WOMAN Ignad 18 O
1026.jpg	1	POV: You're my wife made with mematic
10014.jpg	0	HOOKER MY DICK VIRGIN ME

Table 1: Categories of MEMEs with Examples for Subtask-A

Model	F1 Score
BERT	0.656
Logistic	0.631
SVM	0.584
LSTM	0.651

Table 2: Result of Subtask-A based on different Models

6 Conclusions and Future Work

We have completed the task using various classification algorithms and evaluated the performance of different classification algorithms for Multimedia Automatic Misogyny Identification (MAMI) shared task. Our overall score is 0.656 for subtask-A which were average as compared to other submissions obtained in the Multimedia Automatic Misogyny Identification (MAMI) shared task. We look forward to experimenting with different advance algorithm or neural network models. Also, till now our algorithms works only with text classification. We are looking forward to work in text and image simultaneously for better accuracy and classification. Also, fine tuning the parameters of the algorithm can help in improvement of the overall performance. We shall be exploring these tasks in the coming days.

References

Maria Anzovino, Elisabetta Fersini, and Paolo Rosso. 2018. Automatic identification and classification of misogynistic language on twitter. In *International Conference on Applications of Natural Language to Information Systems*, pages 57–64. Springer.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Elisabetta Fersini, Francesca Gasparini, Giulia Rizzi, Aurora Saibene, Berta Chulvi, Paolo Rosso, Alyssa Lees, and Jeffrey Sorensen. 2022. SemEval-2022 Task 5: Multimedia automatic misogyny identification. In *Proceedings of the 16th International Work-*

shop on Semantic Evaluation (SemEval-2022). Association for Computational Linguistics.

Rachael Fulper, Giovanni Luca Ciampaglia, Emilio Ferrara, Y Ahn, Alessandro Flammini, Filippo Menczer, Bryce Lewis, and Kehontas Rowe. 2014. Misogynistic language on twitter and sexual violence. In *Proceedings of the ACM Web Science Workshop on Computational Approaches to Social Modeling (ChASM)*, pages 57–64.

Iakes Goenaga, Aitziber Atutxa, Koldo Gojenola, Arantza Casillas, Arantza Díaz de Ilarraza, Nerea Ezeiza, Maite Oronoz, Alicia Pérez, and Olatz Perez de Viñaspre. 2018. Automatic misogyny identification using neural networks. In *IberEval@SEPLN*.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

Vipin Kumar and Basant Subba. 2020. A tfidfvectorizer and svm based sentiment analysis framework for text data corpus. In *2020 National Conference on Communications (NCC)*, pages 1–6.

Edward Loper and Steven Bird. 2002. Nltk: The natural language toolkit. *arXiv preprint cs/0205028*.

William S Noble. 2006. What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567.

Endang Wahyu Pamungkas, Valerio Basile, and Viviana Patti. 2020. Misogyny detection in twitter: a multilingual and cross-domain study. *Information Processing & Management*, 57(6):102360.

Abir Rahali, Moulay A Akhloufi, Anne-Marie Therien-Daniel, and Eloi Brassard-Gourdeau. 2021. Automatic misogyny detection in social media platforms using attention-based bidirectional-lstm. pages 2706–2711.

Claude Sammut and Geoffrey I. Webb, editors. 2010. *Logistic Regression*, pages 631–631. Springer US, Boston, MA.

Kalpna Srivastava, Suprakash Chaudhury, PS Bhat, and Samiksha Sahu. 2017. Misogyny, feminism, and sexual harassment. *Industrial psychiatry journal*, 26(2):111.