SOCIAL MEDIA POST IMPACT PREDICTION USING COMPUTER VISION AND NATURAL LANGUAGE PROCESSING

Mihai Minut, Diana Isabela Crainic, Catalin Sumanaru, Ciprian Danis, Ioan Sava, Cristian Simionescu and Adrian Iftene
August 6, 2022

Faculty of Computer Science, Iasi, Romania
# Table of Contents

1. Problem Description

2. Approaches

3. Results

4. Future Work

5. Conclusions
PROBLEM DESCRIPTION
Predicting the outcome of a social media post is valuable to large groups of people and companies.

- Create an application on which the users will make social media post prediction.
- Create a model that will predict the impact.
- Experiment with hybrid deep learning models.
- Measure and evaluate the application and the model.
The task for classifying e-commerce products given a dataset consisting of the image of the products, their descriptive text, and their title. There was proven that better results are given by the hybrid model. [1]

The problem of detecting if a social media post is offensive. The results have shown that the hybrid solutions obtained better accuracy when comparing them to the ones consisting only in an NLP model. [2]
DIFFICULTIES

- complex SOA that incorporates model updates and their learning
- scarce sources of datasets with texts and image
- long training times for the hybrid model
- long and more complex data processing step (images + text)
- long experimental, testing and validation phase for the model
APPROACHES
The current solution proposes a social media post impact predictor. Given the desired image, the neural network must predict the expected number of “likes” or reactions that the post will generate.

The application was extended to also use the description to predict the impact of a post with greater accuracy.
- Micro-services-based architecture
- Account Management Service,
- Prediction Service,
- Predictions Management Service.
- AWS cloud infrastructure.
· Twitter API for data importing
· Different fields for posts
· Each field with word dictionary for searching
· Import data locally by saving metadata, text and image link
· Remove unwanted characters from text
· Download images separately due to high amount of time
· Filter perfect duplicates
· Filter text and image by quality
· Sanitize the data and standardise
It was built a hybrid model ensemble composed of two sub-models: (1) an **NLP model** for processing the description text of the tweet along with (2) a **Computer Vision model** for the image.

The final prediction result was obtained by averaging both previous results.
The training process involved fine-tuning two pretrained models: bert-base-uncased version for the NLP task and Resnet18 for the computer vision one.

**Figure:** The architecture of the Fine tuned BERT base classifier[3]

**Figure:** The architecture of the ResNet 18 model[4]
USE CASES

Beware the pack of cuteness!

Inputs

Model

Prediction on the impact 7 days from now

8th class of impact 2959 - 17284 likes

Cute animals is a currently ongoing trend
Finding the best combination of image and text for a tweet:
contains the image, text, number of likes and some other metrics of 120,000 tweets

- 18 words on average
- 947.7 pixels - average image width
- 675.4 pixels - average image height
- 250GB
- 2-3 days of import
- dimension standard - from 512 down to 16 pixels images (other dimension scaled)
The models achieved on average 10% accuracy scores on the test data at the end of training. The results show that the testing accuracy is pretty low, making the hybrid model build comparable to taking a random guess.
Further model generations should be trained and tested on other variations of the dataset featuring lower feature dimensions and as well as more complex techniques in terms of binding the hybrid model layers together.
FUTURE WORK
APPLICATION & ARCHITECTURE

- Adapt the architecture to also collect data from the users and use it for model training
- Allow the user to select from the best prediction combinations
- Provide the user with a post prediction history search feature
- Overall improvements to the infrastructure for efficiency and stability
A method to provide the dataset in different formats and sizes for the images and for the text.

Training on multiple formats for the data and as well as for much longer times.

New techniques for each model component could be used and refined to bring out a better accuracy.

More experiments on the model architecture

Test on more image size standards and text’s embeddings size
CONCLUSIONS
CONCLUSIONS

- Social media posts impact prediction
- Twitter as target platform
- Service oriented architecture
- Web Client
- Account functionality integrated with Twitter accounts
- Hybrid model utilizing both image and text as input
- Big dataset of tweets consisting in multiple data types, formats and standards
REFERENCES


 tensor.reloaded@gmail.com
Tensor-Reloaded
tensor-reloaded.com
@TensorReloaded
Tensor Reloaded
tensor-reloaded
Tensor Reloaded