

An Assessment of ML-based Sentiment Analysis for Intelligent Web Filtering

Nearchos Paspallis¹, Panayiotis Panayiotou^{1,2}

¹ *University of Central Lancashire (Cyprus) – UCLan Cyprus*

² *Complytek*

Outline

- Background and Motivation
 - Related Work
- Design and Implementation
 - Sentiment Analysis
- Evaluation
 - Data Collection, Analysis, Assessment, Findings
- Conclusions



Background and motivation

- Problem
 - Negative web media content can have an impact on the users' mood and wellbeing
 - Example: Covid-19 and doomscrolling
- Our approach
 - A web-browser plugin to support users by automatically providing clues

Related Work

- AI and ML
 - Numerous applications: autonomous driving/flying
 - Automatic code generation
 - Advent of LLMs gave rise to chatbots, summarization tools, etc.
- Specific to Web filtering
 - Internet provider filtering controls
 - Web browser extensions
 - Athiral et al. [2017] “...filter profanity, bad words and words with double meaning”
 - Noble and Sridhar [2021] – Detox browser is a web filtering tool aiming to “improve users’ wellbeing”

Design and Implementation

- Design
 - Front-end: developed in jQuery provides a control and monitoring interface
 - Back-end: developed in Python and Flask provides the functionality to analyse content and identify its sentiment (positive, neutral, negative)

Sentiment Analysis

- Datasets
 - **Sentiment140 Dataset** (1.6M tweets, very popular)
 - **Twitter Sentiment Analysis Dataset** (32K tweets, but limited scope to hate speech detection)
 - **Twitter and Reddit Sentimental analysis Dataset** (200K tweets and reddit posts, but limited to opinions on government elections)
- Selection and processing
 - Selected Sentiment140 Dataset
 - Cleaned and prepared for assessment of algorithms

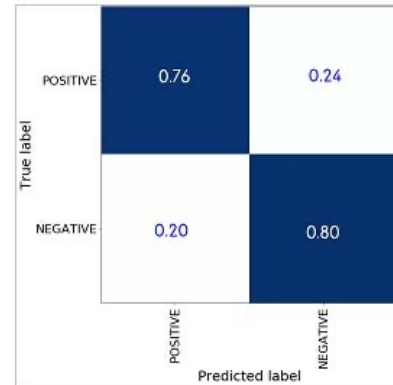
Sentiment Analysis

- Algorithms

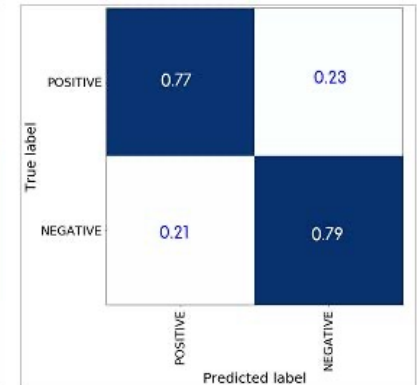
- **Bernoulli Naive Bayes (78%)**
- **Linear Support Vector Machine (Linear SVC) (78%)**
- **Logistic Regression (79%)**
- **Multinomial Naive Bayes (78%)**

- Selection

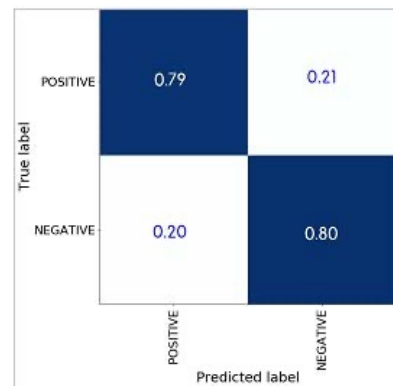
- Similar accuracy (78-79%) and false-neg/false-pos (~20%)
- Also, similar computational performance
- Selected Logistic Regression



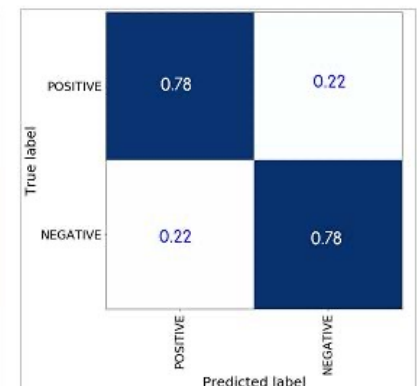
Bernoulli Naive Bayes Model



Linear SVC



Logistic Regression



Multinomial Naive Bayes

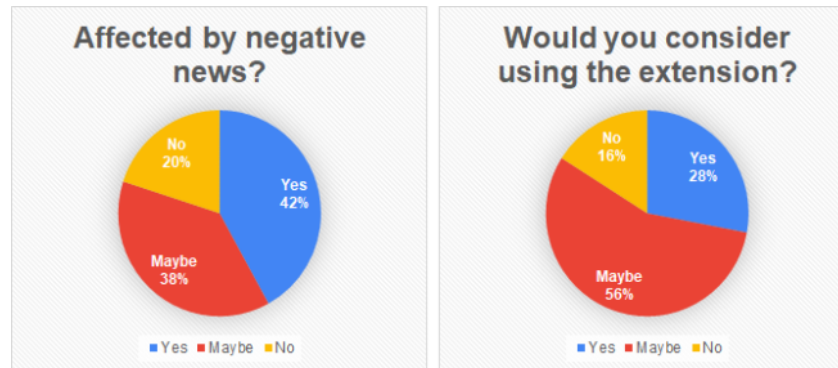


Evaluation

- Data Collection
 - Participants (convenience sampling)
 - Demographics
- Collected Data
 - One-on-one sessions
 - Use the plugin to browse the test website until articles were labeled as negative

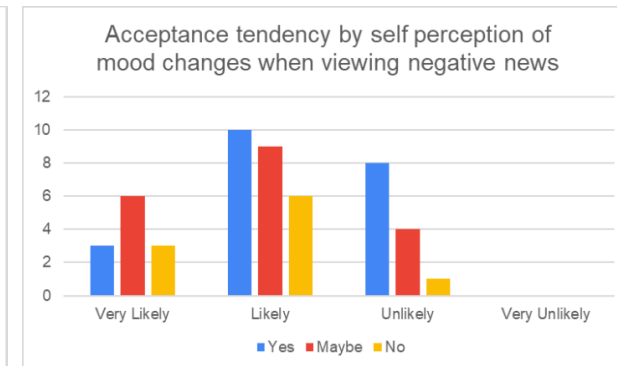
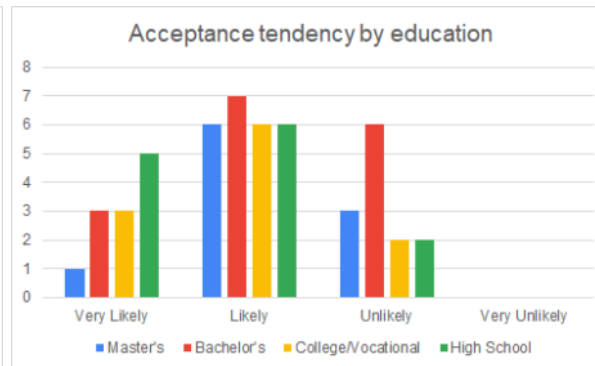
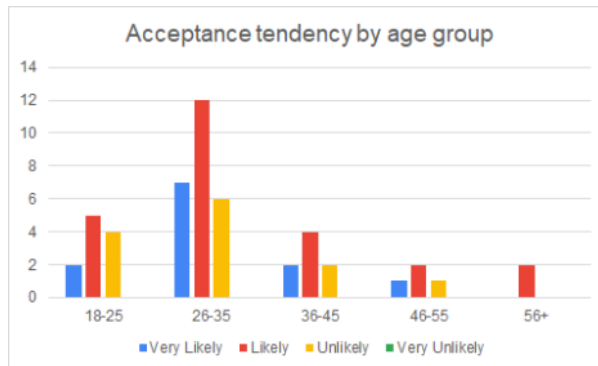
Evaluation

- Assessing the need for the web browser extension



Evaluation

- Assessing the acceptance of the web browser extension
 - By age group
 - By education level
 - By self-perception of mood changes as a result of negative news



Conclusions

- Scope
 - Motivation and proposed solution
 - Selection of data and algorithm
- Evaluation
 - Welcomed by users
 - Acceptance not affected by age or level of education
 - More impact by whether users self assessed themselves as 'affected by negative news'
- Future work
 - Enhance browser extension with feedback (why that label?) and ability to provide their own assessment (do they agree with the assigned label?)

Thank you!



Dr Nearchos Paspallis

NPaspallis@uclan.ac.uk



Panayiotis Panayiotou

p.panayiotou@i-spiral.com