



**INFORMATICS
INSTITUTE OF
TECHNOLOGY**

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

**COMWIR: COMparative Weakness IdentifieR -
identifying Weaknesses of Products from Comparative Reviews**

A Dissertation by

Miss. Noor Amna Hilmy

Supervised by

Mr. Nipuna Senanayake

Submitted in partial fulfilment of the requirements for the BEng in Software
Engineering degree at the University of Westminister.

July 2022

Abstract

A Comparative Opinion evaluates one entity, while keeping the other entities as a reference point for the comparison. Therefore, analysis of such comparative opinions is beneficial for businesses to mine customer requirements and utilize them to enhance the quality of their goods and services than that of their rivals. It also assists businesses in analysing the strengths and weaknesses of their goods while comparing it with that of their rivals. Comparative opinions are different to regular opinions as they provide more precise information through comparison in contrast to vague information provided through direct opinions.

Opinion Mining or Sentiment Analysis is a field in Natural Language Processing (NLP) that has got a lot of attention in recent times. Apart from expressing one's opinions about an entity or aspect as positive or negative, a person can also express opinions by comparing similar and common entities or aspects. Comparative Opinion Mining is a sub-field in Opinion Mining which identifies and extracts information from comparative sentences. Prior works in Comparative Opinion mining were mainly focused on the application of traditional Machine Learning Techniques to mine comparative opinions on limited datasets of comparative product reviews. In contrast, this work is aimed at identifying the comparative opinion to identify the pros and cons of two competing products and extract the compared aspect from simple comparative sentences.

The proposed method consists two modules; a Novel Neural Network to predict the comparative sentiment of a given comparative review a Novel Keyword Extraction approach to identify the common compared aspect with the use of 2 existing keyword extraction libraries. The proposed Sentiment Classifier is a custom CNN and has an accuracy of 92.194% on the hold-out data. Precision, Recall and F1 measures of the classifier on the same data were 80.745%, 82.803% and 81.761% respectively. In addition, the calculated AUC and Average Precision scores were 96.225% and 89.442% respectively. Finally, the author calculated a combined similarity score using ROGUE-1, ROGUE-L and Spacy's similarity scores to determine the most probable comparative aspect among the sets of keywords extracted using YAKE and keywords generated using KeyBERT.

Keywords:

Comparative Opinions, Sentiment Analysis, Product Reviews, Supervised Learning, Neural Networks, Keyword Extraction, Unsupervised Learning, Keyword Extraction, Decision Support Systems, Comparative Weaknesses, Competitor Analysis