# Recommender Systems

Sam Robbins
*Department of Computer Science*
*Durham University*
Durham, United Kingdom

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

*A. Domain of application*

*B. Related work review*

*C. Purpose/aim*

II. Methods

*A. Data description*

*B. Data preparation and feature selection*

*C. Hybrid scheme*

*D. Recommendation techniques/algorithms*

*E. Evaluation methods*

III. Implementation

*A. Input interface*

*B. Recommendation algorithm*

*C. Output interface*

IV. Evaluation results

*A. Comparison against baseline implementation*

*B. Comparison against hybrid recommenders in related studies*

*C. Ethical Issues*

V. Conclusion

*A. Limitations*

*B. Further developments*

References

The data is from the Yelp dataset [1], this contains a list of the businesses on Yelp, detailing which categories they fall under. It also contains all the reviews, including a written review along with a range of scores. This also now contains a COVID-19 dataset, which lists the changes businesses are making due to COVID-19. There are around 8 million reviews in the review dataset.

In order to make the data easier to process, along with increasing the accuracy of predictions, it is needed to reduce the number of reviews. First I selected just restaurants, but as this is a large part of Yelp, it only reduced the dataset to 5 million reviews. So I have reduced it to just Italian restaurants, shrinking the dataset to around 470k reviews. In addition, limiting the timescale reduces the impact of tastes changing over time, so I have only included reviews after 2016. This reduced the number of reviews to around 270k, which is enough to give a good sample, but small enough to be practical to process.

[1] Yelp Dataset. https://www.yelp.com/dataset