

## Tracking a Company's Performance – a Natural Language Processing approach

Ethan Chew<sup>1</sup>

<sup>1</sup>City, University of London, <sup>2</sup>Kainos

### Background and Motivation

With digitisation and globalisation, choices have never been more abundant. Organisations are threatened by emerging players and thus, it is in a company's best interest to continuously improve upon their capabilities to stay one step ahead of others.

This study aimed to design and build an end-to-end Natural Language Processing (NLP) pipeline to obtain insights on performing and trailing capabilities (relatively) through customer reviews - a publicly-available and information rich dataset. Through our pipeline we looked to tag reviews autonomously from a taxonomy of topics.

### Data and Methods

#### Data Source:

- We looked at reviews surrounding UK's major banks – Barclays, Monzo, and HSBC for example.
- 300,000 reviews were scraped from TrustPilot in total, including supplementary data like review ratings (1-5 stars) and bank replies.

#### Methods:

#### Feature Extraction:

- Employed a series of unsupervised learning techniques to group and derive topics for similar reviews. Collectively, these form a taxonomy of topics specific to our domain.
- As seen in Figure 1, similar reviews are clustered together and projected onto a low-order representation of the semantic space.
- Topic modelling then provided a broad understanding of each cluster.

#### Classification:

- Using topics as labels, developed a multi-label classification model which tags reviews with multiple topics.
- A set of Binary Random Forest classifiers are trained and applied. Existing transformer-based LLMs were subsequently fine-tuned for multi-label classification through transfer-learning.

### Key Findings

1. Conventional banks have a much smaller presence online (on TrustPilot) as compared to challenger banks – challenger banks boast 15 times more reviews and reply to 80%, versus 0%, of their unsatisfactory reviews (1-3 stars).
2. Customer service and account issues are the main challenges faced by all banks, collectively making up close to 50% of all unsatisfactory reviews.
3. As an example of internal monitoring, Revolut's app interface takes higher priority for improvements compared to its refund capabilities - each making up 3.3% and 1.6% of unsatisfactory reviews respectively
4. As an example of benchmarking, Wise saw its app-interface lagging behind Revolut's with unsatisfactory reviews (regarding app-interface) leading up to 2023, from 2022, increasing by 140% and decreasing by 42% respectively.

### Value of the research

This study provided a practical approach to process unstructured reviews into actionable insights. The learnings from this study are applicable across several domains and allow for companies to understand the performance of their functions, as well as their competitors. Through it, companies can adopt best practices, capitalise upon missed opportunities, proactively identify, prioritise areas for further development, and continue to maintain a marginal edge.

Figure 1: Visualisation of reviews in the semantic space

