**Report**

| Abstract |  | ZHANG XIAOFANZHU YUEJIE |
| --- | --- | --- |
| Introduction | Research Goal | HU YUZHANG XIAOFAN |
| Why is the Topic Important |
| Related Work | Literature Review | HU YUZHANG XIAOFAN |
| Research Gap |
| Sample data | Data Extraction | WU CHUAN |
| Data Preprocessing | ZHU YUEJIE(3.2.1. and 3.2.3.)WU CHUAN(3.2.2.) |
| Overall overview of end-use variables | ZHU YUEJIE |
| Exploratory Data Analysis | Quantitative data | ZHU YUEJIEZHANG XIAOFAN |
| Text data | WU CHUAN |
| Model Building and Evaluation | Quantitative data | ZHU YUEJIEZHANG XIAOFAN |
| Text data | WU CHUAN |
| Conclusion and Future Recommendations | Conclusion | WU CHUAN HU YU  ZHU YUEJIE |
| Limitations and future work | ZHU YUEJIEWU CHUAN |
| Article modification |  | HU YU |

**Code**

| Data pre-processing.ipynb | ZHU YUEJIE  HU YU |
| --- | --- |
| Quantitative\_data\_Users.ipynb.ipynb | Visualization+Original model: ZHANG XIAOFAN  Unbalanced data handling and model evaluation: ZHU YUEJIE |
| Quantitative\_data\_Published\_content.ipynb | ZHU YUEJIE |
| Time series prove not to be available.ipynb | ZHU YUEJIE |
| LSTM.ipynb | ZHU YUEJIE  WU CHUAN |
| MultinomialNB.ipynb | ZHU YUEJIE  WU CHUAN |
| Wordcloud\_map\_and\_distribution.ipynb | ZHU YUEJIE |
| LR.ipynb | WU CHUAN |
| GRU.ipynb | WU CHUAN |
| LDA Visulization.ipynb | WU CHUAN |