

Counterfeit Review Detection — One Pager

****Author:**** Yugal Soni · MSc Business Analytics & Decision Science (University of Leeds)

Objective

Detect ****fake product reviews**** and flag ****high risk entities**** (sellers/categories) prone to counterfeit activity using a combined ****classification + risk scoring**** approach.

Data

- `trustpilot_reviews.csv`, `sample_trustpilot.csv`: product reviews scraped from Trustpilot (text, rating).
- `fake_reviews_dataset.csv`: labelled reviews used to train/evaluate a fake review classifier.

Methods

- ****Classifier:**** TF-IDF features + ****Linear SVM**** baseline (benchmarked vs Logistic Regression / RandomForest).
- ****Metrics:**** Accuracy, Precision, Recall, F1, ****ROC-AUC****; threshold sensitivity analysis.
- ****Risk Scoring (CRS):**** weighted combination of indicators:
 - ****Fake review ratio****, ****rating polarisation****, ****review burstiness****, ****sentiment rating gap****, ****lexical redundancy**** (duplication).
 - Scaled per category (z-scores), mapped to ****risk levels**** (Low/Medium/High/****Very High****).
- ****Unsupervised checks:**** ****K-Means**** (segment entities by risk profile) and ****Isolation Forest**** (anomaly detection).

Key Results

- ****Model performance (test set):**** Accuracy ****≈ 90%****, Precision ****≈ 90.7%****, Recall ****≈ 89.5%****, ****ROC-AUC ≈ 96.6%****.
- ****Entities analysed:**** 63. ****Very High risk ≈ 14%****, ****High risk ≈ 43%**** of entities; remainder Medium/Low.
- ****Top flagged examples**** (for illustration): Target (Books), Walmart (Beauty), Lowe's (Sports), HomeDepot (Electronics), Amazon (Automotive).

Business Impact

- ****Trust & Safety:**** Faster triage of suspect listings; prioritise ****Very High**** and consensus-flagged entities for investigation.
- ****Marketplace Ops:**** Monitor ****review burstiness**** and ****fake ratio**** spikes; enforce stricter seller audits in hot-spot categories.
- ****Analytics:**** Monthly dashboards to track risk trends and the effect of interventions.

Reproduce

```
``bash
pip install -r requirements.txt
python code/01_classifier.py
python code/02_counterfeit_risk.py
...`
```

Outputs are written to `/outputs` and mirrored in the README.