



PSX PNR: Pakistan Stock eXchange Personalized News Recommendation

ADAM Muhammad Safi Ullah, CERESA Thomas

Section I - Problem Statement

- Too much news makes it hard for investors to find what truly matters to them. **Time Consuming**
- Most platforms show general market news, not **personalized** stock-specific updates.
- A smart system is needed to **recommend relevant stock news** based on user interests especially for PSX investors..



Section I – Project Objectives (Motivation)

- ❑ Deliver personalized stock-market news tailored to each user's interests and sectors (Banking/Construction/Technology).
- ❑ Reduce information overload by filtering and ranking only the most relevant articles.
- ❑ Use NLP embeddings (Word2Vec/BERT) to understand the semantic meaning of financial news and effect on stocks.
- ❑ Provide an easy interface where users can quickly access top-K relevant market updates daily under 20 minutes.

Section II - Approach and Datasets

Google Scholar

personalized news recommendation



Articles

About 405 000 results (0,09 sec)

Any time

Since 2026

Since 2025

Since 2022

Custom range...

Personalized news recommendation: Methods and challenges

[C Wu](#), [F Wu](#), [Y Huang](#), [X Xie](#) - ACM Transactions on Information Systems, 2023 - dl.acm.org

... master the advances in **personalized news recommendation**, in this ... of **personalized news recommendation**. Instead of following the conventional taxonomy of **news recommendation** ...

☆ Save Cite Cited by 292 Related articles All 4 versions

[PDF] acm.org

Full View

- ① Confirms Our Proposed Pipeline Design
- ② Justifies Using NLP Embeddings for Semantic News Text
- ③ Explains Why User Profiling Matters - Use of Synthetic Users due to time and resource constraints

Section II – Datasets


- **Pakistan News Headlines (Kaggle):** Provides a large general news corpus to understand Pakistani news language and context.
- **Pakistan Stock Market Data 2017–2025 (Kaggle):** Supplies historical PSX price trends to analyze market movements and relate major news events to stock activity.
- **CNH-PSX Categorized Financial News (Mendeley):** Offers finance-focused, PSX-relevant news with category labels.

[Sign In](#) [Register](#)

Pakistan Stock Market Data 2017–2025

Daily open, high, low, close, and volume data for all companies listed on the Pa

[Data Card](#) [Code \(1\)](#) [Discussion \(0\)](#) [Suggestions \(0\)](#)




About Dataset

This dataset provides historical daily trading data for companies listed on the Pakistan Stock Exchange (PSX). It includes records of stock prices, volumes, and percentage changes across all active and delisted companies over the period January 2017 – October 2025.

The dataset can be used for:

- Financial market trend analysis
- Predictive modeling of stock movements
- Time-series forecasting
- Quantitative finance and algorithmic trading research
- Data visualization and analytics education


 Mendeley Data

CNH-PSX





Published: 9 December 2021 | Version 1 | DOI: 10.17632/mc4s7zvx9c.1
Contributors: [Shazia Usmani](#), [Jawwad Ahmed Shamsi](#), [Uzma Afzal](#)

Description

The categorized financial news has been used significantly in stock prediction. Therefore, PSX-related financial news is categorized into the market, sector, and stock categories. Moreover, sector and stock categories are further filtered to represent the specific sectors and stocks categories. The finalized datasets are stored into two (.csv) files.


[Download All 474 KB](#) 

Files

| | |
|--|--|
|  CNH-PSX_Ver1.csv | 838 KB  |
|  CNH-PSX_Ver2.csv | 1.1 MB  |

Steps to reproduce

- This dataset is prepared for news headlines related to Pakistan Stock Exchange (PSX) and contains news headlines along with publishing date.
- Financial news headlines are scraped from daily "The News International" and properly categorized according to their area of influence in the stock market.
- The news categorization scheme used for categorization is unsupervised.
- News categorization scheme uses strength of NLP along with minimal manual effort.
- Finalized dataset is validated using ANN based classification model.

Usability 

10.00

License

Attribution 4.0 International (CC BY)

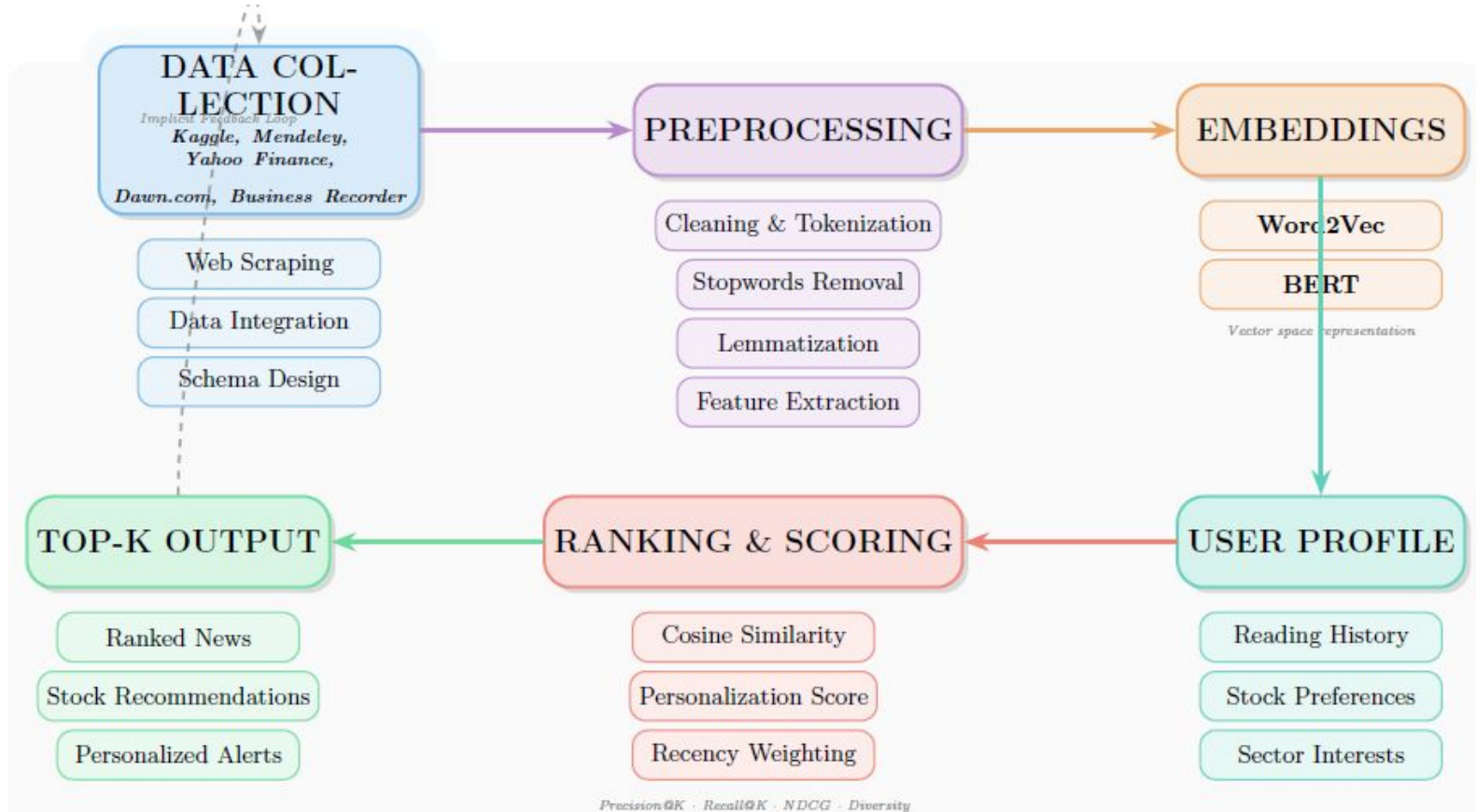
Expected update frequency

Quarterly

Tags

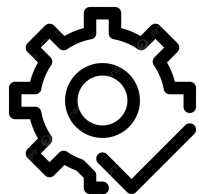
[Finance](#) [Economics](#) [Time Series Analysis](#) [Asia](#)

Section III - System Overview



Section III - Development

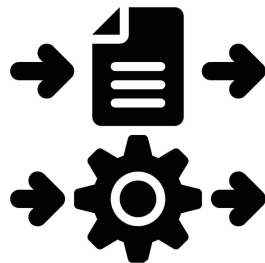
Phase 1: Fixed Corpus (Baseline)



(S)BERT
vs
Word2Vec



Phase 2: Advanced Features (If Time Permits, or in Parallel)



Thank you

