Sentiment Analysis of RBI Monetary Policy Statements

Machine Learning and Pattern Recognition Final Project

Sem 4

The Problem Statement


Home Latest E-Paper Markets Elections Opinion India News Portfolio Technology

Home / Markets / News / Financial stocks feel the pinch of RBI loan stance; Sensex falls 188 points

Financial stocks feel the pinch of RBI loan stance; Sensex falls 188 points

Sensex, Nifty decline intraday but finish with gains for a third week in a row


The Economic Times | MutualFunds

English Edition | Today's ePaper

Home ETPrime Markets News Industry Rise Politics Wealth Mutual Funds Tech Careers Opinion NRI Panache ET TV

Web Stories MF News Analysis Mutual Fund Screener ELSS Learn ETF Best Mutual Funds to Buy NPS Tools MF Recatego

Business News › Mutual Funds › Analysis › Debt fund managers decode RBI policy for investors

Debt fund managers decode RBI policy for investors

ET Online • Last Updated: Oct 06, 2023, 04:06:00 PM IST

Synopsis

"RBI would target Inflation rate of 4% and not 2%-6% range, again signals the policy rates to remain higher for longer until Inflation is projected to come below 4%. RBI's sharp focus on bringing down inflation is positive for markets in the medium-to-long term," says Amit Somani, Senior Fund Manager – Fixed Income, Tata Asset Management.


mint

Markets Premium Money Mutual Fund Industry Companies Techn

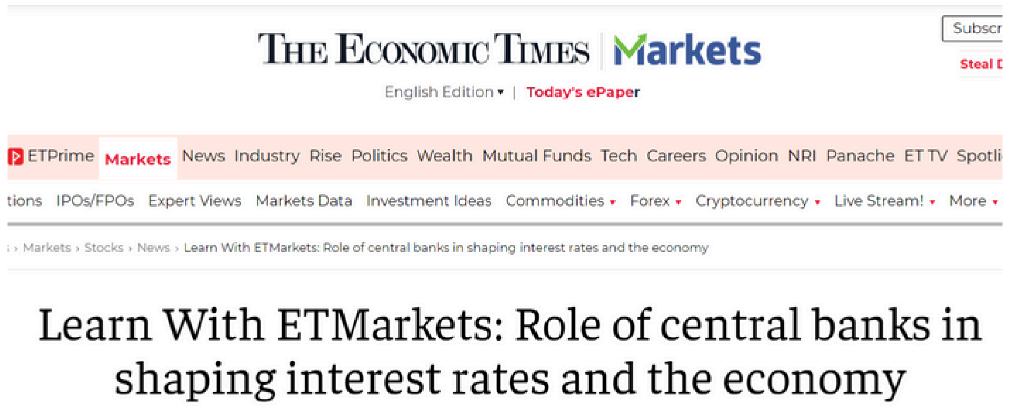
Mint+The Wall Street Journal at ₹3499 Subscribe

Business News / Markets / Stock Markets / Why stock market is down today after...

Why stock market is down today after RBI's announcement to keep interest rate unchanged?

2 min read • 08 Feb 2024, 12:04 PM IST

Asit Manohar


The Economic Times | Markets

English Edition | Today's ePaper

ETPrime Markets News Industry Rise Politics Wealth Mutual Funds Tech Careers Opinion NRI Panache ET TV

IPOs/FPOs Expert Views Markets Data Investment Ideas Commodities Forex Cryptocurrency Live Stream! More

Markets › Stocks › News › Learn With ETMarkets: Role of central banks in shaping interest rates and the economy

Learn With ETMarkets: Role of central banks in shaping interest rates and the economy

Historically, experts opine that RBI communications, especially in the form of monetary policy statements, show an impact on the market. Our project aims to leverage machine learning to quantify these communications in the form of a sentiment score and predict their impact across financial indices. This would equip risk bearing individuals to make better investment decisions in the long run.

The Purpose

Communications in Central Banks: A Perspective

Rakesh Mohan*

I. Introduction

Communication is part of the professional hazards that central bankers face as a routine. The responses as well as measures are mostly ‘measured’. Financial markets watch our pauses, and some punctuation marks in the text of a statement have the distinct possibility of getting transmitted through the movement of a few points in the yield curve. Faced with such grave consequences that can be measured in millions of rupees or dollars, communication in central banking becomes really a very serious matter.

Managing market expectations:

Central banks influence the economy through both direct actions and shaping market expectations via nuanced communication.

Interpreting Policy Stance and Rationale:

Central bank communications provide insights into their policy stance, objectives weighting, and decision rationale.

Assessing Transparency and Credibility:

Analyzing central bank language clarity and consistency gauges institutional credibility and transparency, enhancing policy effectiveness and accountability.

1. Mohan, R., & Ray, P. (December, 2009). Communications in Central Banks: A Perspective.

Literature Review

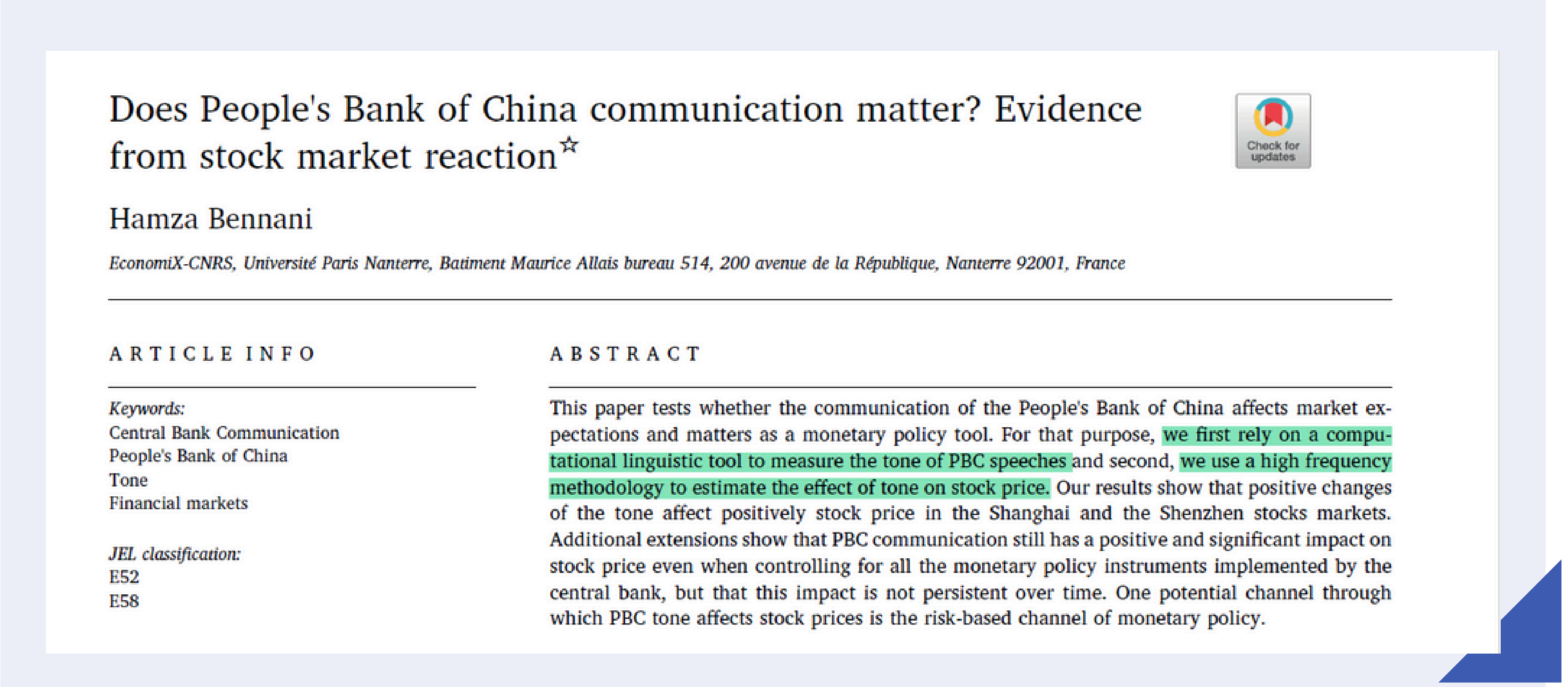
PAPER 1:

Bennani, H. (n.d.). ***Does People's Bank of China communication matter? Evidence from stock market reaction☆***. EconomiX-CNRS, Université Paris Nanterre, Batiment Maurice Allais bureau 514, 200 avenue de la République, Nanterre 92001, France.

Review:

Context: The paper explores the impact of People's Bank of China (PBC) communication tone on the stock market.

Dataset: PBC speeches are analyzed to measure tone and its impact on stock prices using a high-frequency methodology.



Does People's Bank of China communication matter? Evidence from stock market reaction☆

Hamza Bennani

EconomiX-CNRS, Université Paris Nanterre, Batiment Maurice Allais bureau 514, 200 avenue de la République, Nanterre 92001, France

ARTICLE INFO

Keywords:
Central Bank Communication
People's Bank of China
Tone
Financial markets

JEL classification:
E52
E58

ABSTRACT

This paper tests whether the communication of the People's Bank of China affects market expectations and matters as a monetary policy tool. For that purpose, we first rely on a computational linguistic tool to measure the tone of PBC speeches and second, we use a high frequency methodology to estimate the effect of tone on stock price. Our results show that positive changes of the tone affect positively stock price in the Shanghai and the Shenzhen stocks markets. Additional extensions show that PBC communication still has a positive and significant impact on stock price even when controlling for all the monetary policy instruments implemented by the central bank, but that this impact is not persistent over time. One potential channel through which PBC tone affects stock prices is the risk-based channel of monetary policy.

ML Approach: bag-of-words approach combined with the financial dictionary by Loughran and McDonald (2011) to identify negative words quantifying the relative degree of negative or positive tone.

$$\tau_t = 1 - N_t/T_t$$

Control Variables: data related to PBC's policy action, macroeconomic fundamentals, external factors, supply shock, markets' volatility, consumer and business confidence and economic uncertainty.

PAPER 2:

Mathur, A., & Sengupta, R. (April, 2020). *Analysing monetary policy statements of the Reserve Bank of India*. Indira Gandhi Institute of Development Research, Mumbai.

Review:

Analysing monetary policy statements of the Reserve Bank of India

Aakriti Mathur and Rajeswari Sengupta

Email(corresponding author): rajeswari@igidr.ac.in

Abstract

We quantitatively analyse the monetary policy statements of the Reserve Bank of India (RBI) between 1998–2018, across five governor regimes. Using natural language processing tools, we show that there has been a persistent semantic shift in RBI's monetary policy communication since adoption of inflation targeting. We construct measures of linguistic and structural complexity that capture governor-specific trends in communication. RBI's communication is linguistically complex on average, but the length of monetary policy statements has gone down and readability has improved significantly recently. Our results indicate that lengthier statements are linked to higher volatility in equity and currency markets, but not bond markets.

Context: examines the linguistic changes in RBI's MP statements pre and post-inflation targeting (IT) implementation and their impact on financial markets.

Dataset: monetary policy statements of RBI from pre and post IT periods

---

ML Approach:

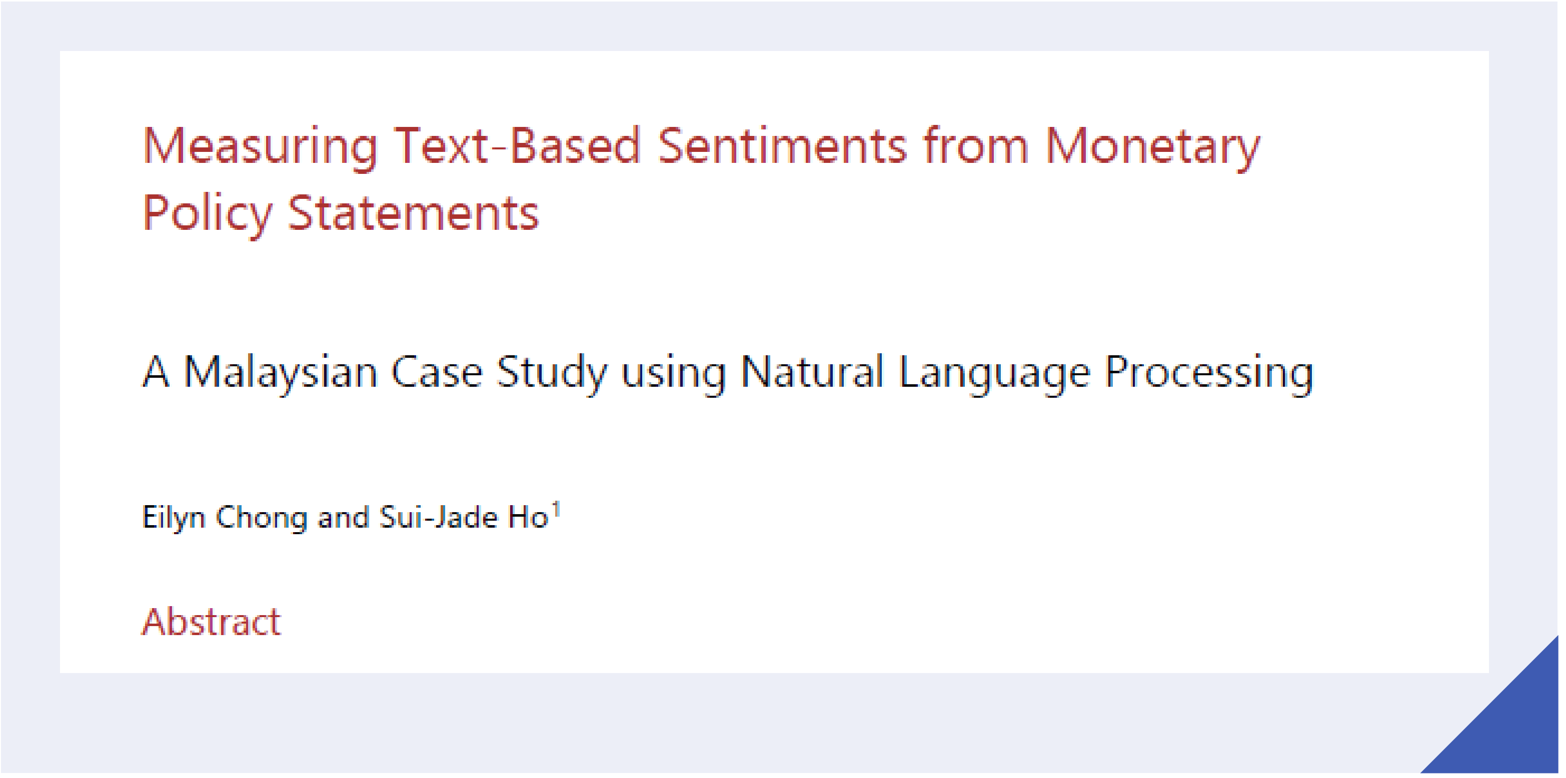
- Employs readability indices like the Farr-Jenkins-Paterson (FJP) index to assess statement complexity.
- Applies ordinary least squares regression to examine the association between linguistic complexity and financial market volatility.

Financial Metric Used: INDIAVIX

PAPER 3:

Chong, E., & Ho, S. (2022). *Measuring text-based sentiments from monetary policy statements – a Malaysian case study using natural language processing.* Central Bank of Malaysia.

Review:


Measuring Text-Based Sentiments from Monetary Policy Statements

A Malaysian Case Study using Natural Language Processing

Eilyn Chong and Sui-Jade Ho[1]

Abstract

Context of the paper: extracting sentiments from MPS of the Central Bank of Malaysia and assessing their impact on financial markets,

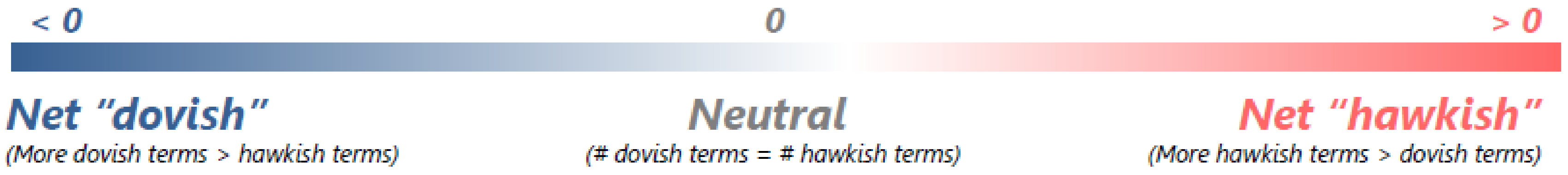
---

Dataset: MPS published by the Central Bank of Malaysia from August 2004 to September 2020

ML Approach:

- Automated content analysis employed to extract sentiments from the MPS texts.
- Three dictionaries method used
- Sentiment derived based on the frequency of words categorized as positive/hawkish, negative/dovish

$$sentiment_t = \frac{\sum positive\ or\ hawkish\ words_t - \sum negative\ or\ dovish\ words_t}{\sum word\ count_t} \text{ for each MPS } t$$



Financial Metric Used: Government Securities (MGS) yields and interest rate swap (IRS) rates within a one-day window

Data Collection and Preprocessing

Data Data Data....

The data includes:

- Bi-monthly minutes of the Monetary Policy Committee (FY 2014-15 to Present)
- The Governer's statement after the quarterly monetary policy review (FY 2013-14 to FY 2005-06)
- Annual Monetary and Credit Policy & its mid-term review (FY 2004-05 to FY 2000-01)
- FOMC minutes of meetings (FY 2000 to Present)

Data was collected by extracting text from PDFs and marking them to a date manually. Basically, we created the dataset!

Steps to clean the data:

01. Removed punctuations and stopwords

02. Converted to all lowercase

03. Performed Lemmatization

MONETARY POLICY

Feb 22, 2024
Minutes of the Monetary Policy Committee
- Minutes of the Monetary Policy Committee Meeting, February 6 to 8, 2024

Feb 09, 2024
Bi-monthly Monetary Policy Statement, 2023-2024
- Edited Transcript of the Reserve Bank of India's Post-Monetary Policy Press Conference: February 8, 2024

Feb 08, 2024
Bi-monthly Monetary Policy Statement, 2023-2024
- RBI Governor's address
- Governor's Statement: February 8, 2024
- Resolution of the Monetary Policy Committee (MPC) February 6 to 8, 2024
  - Full Document
- Statement on Developmental and Regulatory Policies
  - Full Document

Dec 22, 2023
Minutes of the Monetary Policy Committee
- Minutes of the Monetary Policy Committee Meeting, December 6 to 8, 2023

Feature extraction

Dictionary Approach

What is it?

- Assigns pre-defined sentiment scores to words or phrases.
- Utilizes a pre-built dictionary or lexicon.
- Analyzes text by comparing sentiment scores of individual words.
- Determines the overall sentiment of the text based on these scores.

Our Approach

For this method, we use three dictionaries:

- Two are off-the-shelf dictionaries: one was developed by **Loughran & McDonald**
- (2011) (LM hereafter) tailored specifically to finance, and the other was developed by **Correa et al. (2021)**, which is a refinement of the LM dictionary catered to the financial stability context. The features we have extracted are **positive score, negative score, neutral proportion**, and **sentiment score** ((positive-negative)/total)
- We then construct another dictionary that combines both LM and Correa dictionaries and refine it to better fit the monetary policy context.

Readability Approach

What is it?

- Assess the ease of understanding financial or policy statements.
- Farr-Jenkins-Paterson (FJP) index quantifies readability, incorporating factors tow factors - one-syllable words per 100 words and average number of sentences.
- Readability analysis considers grammar and sentence structure.
- Reveals trends in communication clarity, impacting market sentiment understanding.

Our Approach

- Using the formula for FJP index, we calculated a **readability score** and used it as one of our features.

```
FJP = 1.599*(num_monosyllabic_words/100) - 1.015*(num_words/num_sent) - 31.517
```

3. Mathur, A., & Sengupta, R. (April 2020). Analyzing monetary policy statements of the Reserve Bank of India. Indira Gandhi Institute of Development Research, Mumbai.

Results

Date	Cleaned_Text	Sentiment_Score	Positive_Score	Negative_Score	Neutral_Proportion	Readability_Score	Normalized_readability
010414	first monetary policy statement raghuram rajan governor part monetary poli	0.002581756	56	50	0.083476764	-2379.75964	0.921683348
010818	section 45zl reserve bank india act 1934 twelfth meeting monetary policy con	0.002967359	139	126	0.129650765	-4455.89397	0.803291682
011215	bi monetary policy statement 16 raghuram rajan governor monetary liquidity	-0.001640689	43	45	0.114027892	-1263.62124	0.985331198
020216	monetary liquidity measures basis assessment current evolving macroeconom	0.008821171	67	56	0.134723336	-1291.40164	0.983747019
020615	press release department communication central office mumbai 91 22 2266 0	-0.018143754	36	62	0.102581996	-1479.00838	0.973048735
020817	press release department communication central office mumbai 22610835 91	0.000231214	177	176	0.10867052	-4401.5644	0.806389829
021110	reserve bank india second quarter review monetary policy including review de	0.010542378	195	119	0.081009849	-7318.83065	0.640032572
021214	1 press release department communication central office umbai 91 22 2266 0	0.002204262	59	56	0.116091109	-1406.93575	0.97715868
030215	press release department communication central office 91 22 2266 0502 91 2	0.003577107	76	67	0.104531002	-2574.36781	0.910585808
030511	rese rve bank ind ia monetary policy statement subbarao governor may 3 201	0.008755097	235	162	0.093907412	-8457.09045	0.575123249
030513	ireserve bank india monetary policy statement subbarao governor may 3 201:	0.001288475	208	199	0.105082319	-7091.82243	0.652977728
030614	press release department communication central office 91 22 2266 0502 91 2	0.007253886	32	25	0.098445596	-1006.38688	1
040419	press release department communication central office mumbai 022 2261 08	-0.007824143	210	252	0.117175857	-5453.0139	0.746430867
040522	may 18 202 2 minutes monetary policy committee meeting may 2 4 2022 sect	-0.008476971	112	142	0.132805877	-3607.14829	0.851691448
040621	1 press release department communication central office mumbai 91 22 2266	0.018501388	201	121	0.132516189	-4399.51005	0.806506978
040815	press release department communication central office mumbai 91 22 2266 0	-0.006228374	51	60	0.135640138	-1491.44422	0.97233958
041016	press release department communication central office mumbai 91 22 2266 0	0.014720812	84	55	0.104568528	-2022.56032	0.942052658
041017	press release department communication central office mumbai 022 2261 08	0.000700607	183	180	0.114198972	-4357.39173	0.808908778
041019	press release department communication central office mumbai 22660502 re	-0.005628518	197	227	0.113320826	-5415.75508	0.748555554
041220	1 press release department communication central office mumbai 91 22 2266	0.028876582	296	150	0.126780063	-5140.13952	0.764272545
050221	1 press release department communication central office mumbai 22660502	0.028565338	243	109	0.119377531	-4770.83179	0.785332337
050416	press release department communication central office mumbai 91 22 2266 0	0.006413832	104	81	0.097322922	-3655.89264	0.848911799
050418	1 press release department communication central office mumbai 91 22 2266	0.013979497	197	137	0.117893756	-4366.50238	0.808389242
050814	press release department communication central office 91 22 2266 0502 91 2	0.024590164	37	13	0.109631148	-1017.2001	0.999383376
050822	august 19 2022 minutes monetary policy committee meeting august 3 5 2022	0.005215647	178	152	0.132798395	-5066.15572	0.768491475
051018	1 press release department communication central office marg mumbai 91 22	0.006438721	184	155	0.124777975	-4580.89887	0.796163272
051218	press release department communication central office mumbai 91 22 22660	0.004306632	202	182	0.127260982	-4722.39125	0.788094661
051219	press release department f communication central office mumbai reserve bar	0.004010939	216	194	0.109024613	-5571.11331	0.739696243
060220	press release department communication central office mumbai 22660502 re	0.006252112	224	187	0.1176073	-6008.94535	0.714728848
060417	press release department communication central office mumbai 91 22 2266 0	0.010848929	138	98	0.098454028	-3756.37691	0.843181678
060423	april 20 2023 minutes monetary policy committee meeting april 3 5 6 2023 se	0.007347876	159	127	0.111825488	-4429.91971	0.804772866
060618	press release department communication central office mumbai 022 2261 08	0.026699629	200	92	0.120889988	-4116.70881	0.822633733

Word Embedding
Doc2vec

Variable length text documents → Vectors with fixed length dimensions

- Unsupervised algorithm used to generate distributed representations of documents.

PV-DBOW (Distributed Bag of Words): In this model, the document vector is trained to predict words randomly sampled from the document. It does not consider the context words but preserves the semantics of the document.

Parameters for doc2vec: A vector size of 120, a window-size of 20, trained over 25 epochs, and using the distributed memory algorithm

“The meaning of words lies in their use” - Wittgenstein

What financial data do we have??

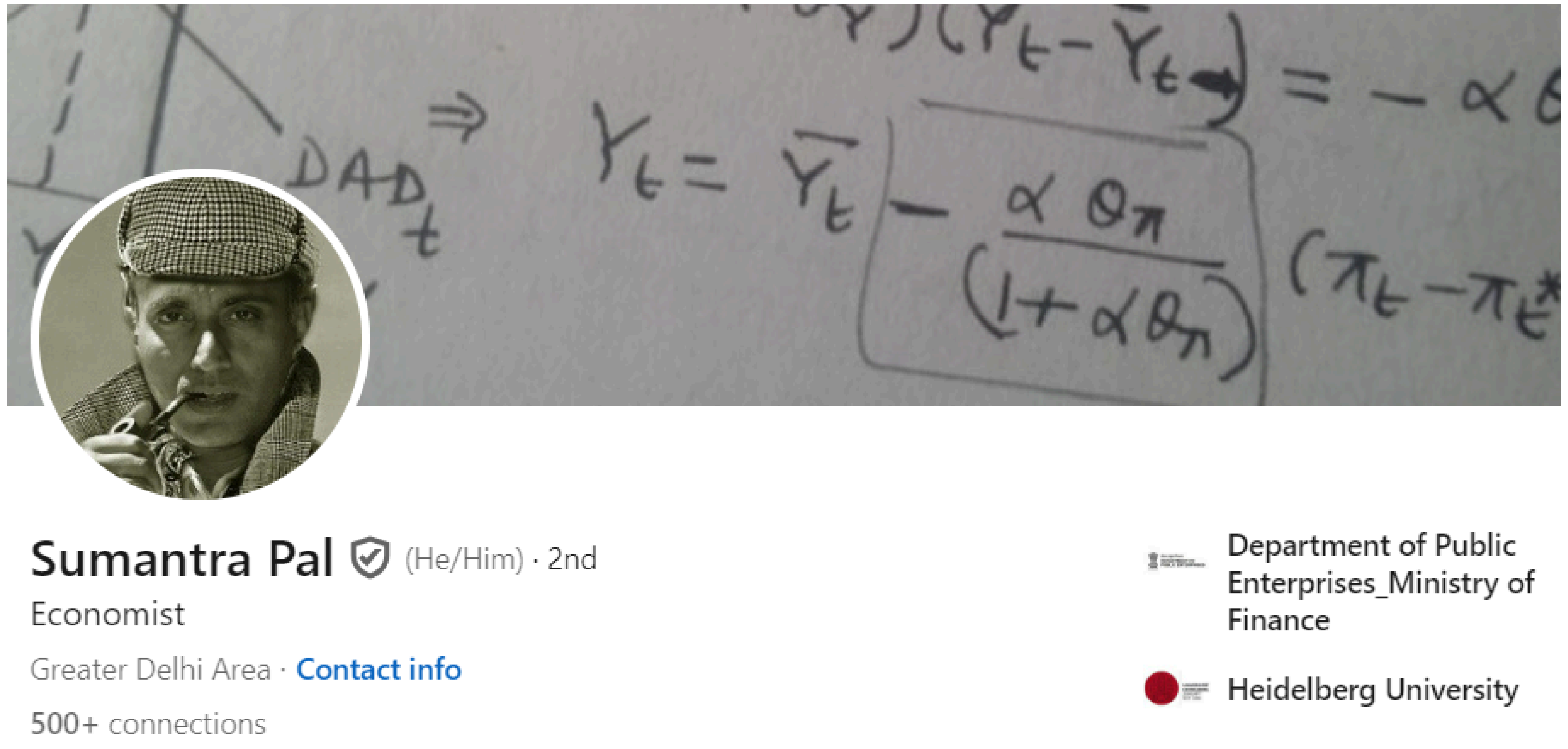
NIFTY50 weekly delta

Sourced from Jefferies FinHub Refinitiv Eikon

NIFTY 50 Value	NIFTY 50 Value After 14 Days	NIFTY Percentage Change
6721.049805	6733.100098	0.179291827
11346.2002	11385.0498	0.342401938
7954.899902	7700.899902	-3.19300058
7455.549805	7048.25	-5.463041833
8236.450195	8047.299805	-2.296503787
10081.5	9897.299805	-1.827110998
6119	5988.700195	-2.129429722
8524.700195	8067.600098	-5.36206655
8756.549805	8869.099609	1.285321356
4494.649902	5068.950195	12.77741994
5565.25	5438.950195	-2.269436318
5944	6187.299805	4.093199944
7415.850098	7631.700195	2.910658857
4494.649902	5068.950195	12.77741994
11598	11752.7998	1.334711198

Stakeholder engagement

In conversation with Dr. Sumantra Pal



$$\boxed{\frac{12}{52} \times 2 = 0.46}$$

We had a conversation with Dr. Sumantra Pal, a member Indian Economic Service, who visited the Plaksha Campus recently.
From this conversation, we understood that there would be a considerable amount of noise in our target variable.

To offset this, he suggested to subtract all values by the benchmark return of NIFTY50, which is 12% p.a.

Data Augmentation

- **Back translation:** We used this technique to increase the size of our train set by 2x. Each entry of data was first translated into French and then back to English to give a new and unique sentiment score.

1	Date	
2	10414	press release department communication central office 91 22 2266 0502 91 22 22660358 reserve bank india website email helpdoc april C
3	10414	press release 1 communication department central office 91 22 2266 0502 91 22 22660358 reserve bank india website website email help
4	10818	press release department communication central office mumbai 91 22 22660358 reserve bank india website email helpdoc august 16 201
5	10818	press release communication department central office mumbai 91 22 22660358 reserve bank india website email helpdoc august 16 201
6	11215	press release department communication central office mumbai 91 22 2266 0502 91 22 22660358 reserve bank india website email helpd
7	11215	press release communication department central office mumbai 91 22 2266 0502 91 22 22660358 reserve bank india website email helpd
8	20216	press release department communication central office mumbai 91 22 2266 0502 91 22 22660358 reserve bank india website email helpd
9	20216	press release communication department central office mumbai 91 22 2266 0502 91 22 22660358 reserve bank india website website ema
10	20615	press release department communication central office mumbai 91 22 2266 0502 91 22 22660358 reserve bank india website email helpd
11	20615	press release communication department central office mumbai 91 22 2266 0502 91 22 22660358 reserve bank india website website ema
12	20817	press release department communication central office mumbai 22610835 91 22 22660358 reserve bank india website email helpdoc aug
13	20817	press release communication department central office mumbai 91 22 22660358 reserve bank india website email helpdoc august 16 201

ML Methodology

- **LSTM:** Long Short-Term Memory

Suitable for long textual data, and time series data, which fits the bill well for our analysis of the reserve bank statements.

Two approaches:

- Breaking the document into clusters of 450 words each
- Maintaining document integrity to build context

LSTM

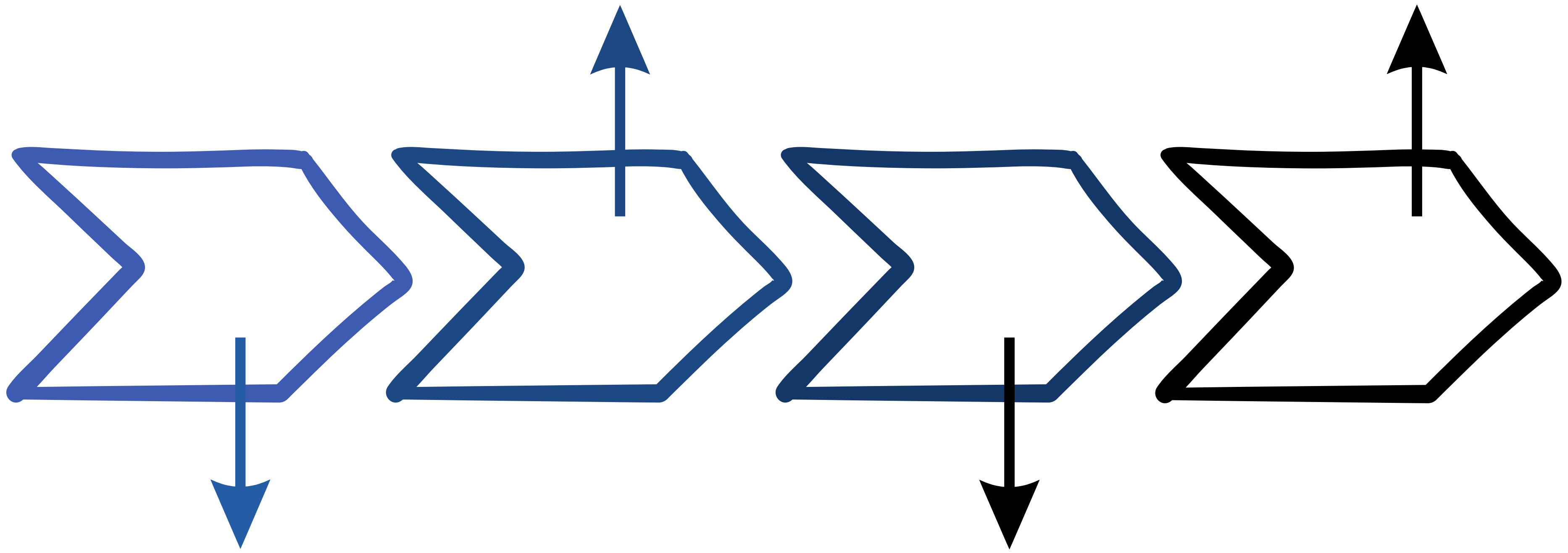
What is it?

- a type of recurrent neural network (RNN) architecture, designed to overcome the limitations of traditional RNNs in capturing long-term dependencies in sequential data.

Our Architecture:

Dropout of 0.4

Dense Layer:
3 Output neurons



Input layer:
128 neurons

Custom Attention
Layer

Hyperparameter Tuning

=

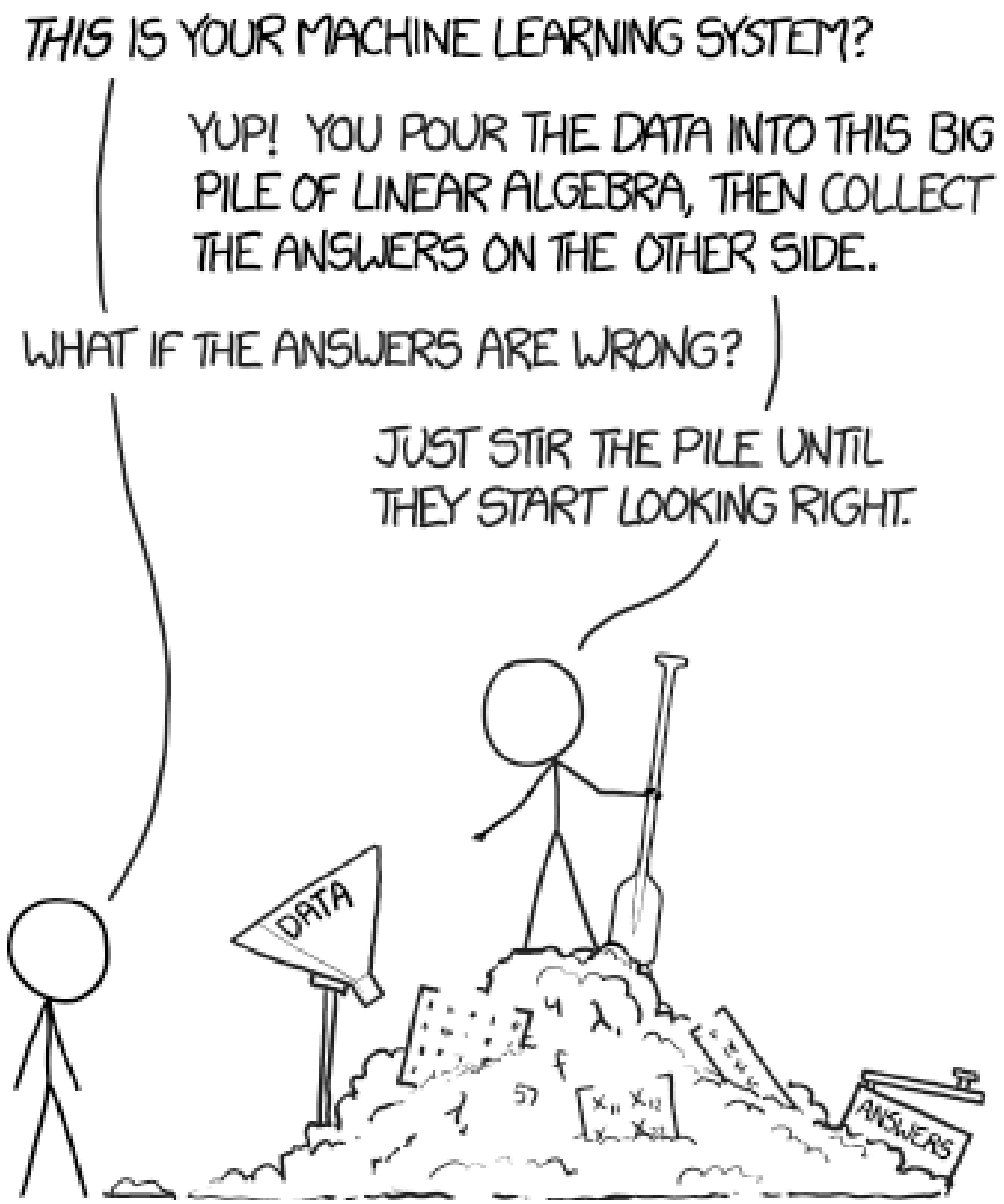
- Batch size of 7 and 8 - experimentally found that lower values yielded better results
- Normalizing using MinMaxScaler()
- Trained for 20 epochs when using RMSProp
- Trained for 15 epochs with Adam - this gave us better results

- We chose the “sparse_categorical_crossentropy” loss function because our target variable was integer type, and this loss function directly calculates the loss between true labels and the predicted probabilities

- We use a 'softmax' activation function as it calculates the probabilities of our categories, which is useful for our purpose.

- We've also added dropouts of 0.4 in the model to help with the overfitting problem.

Performance Metrics

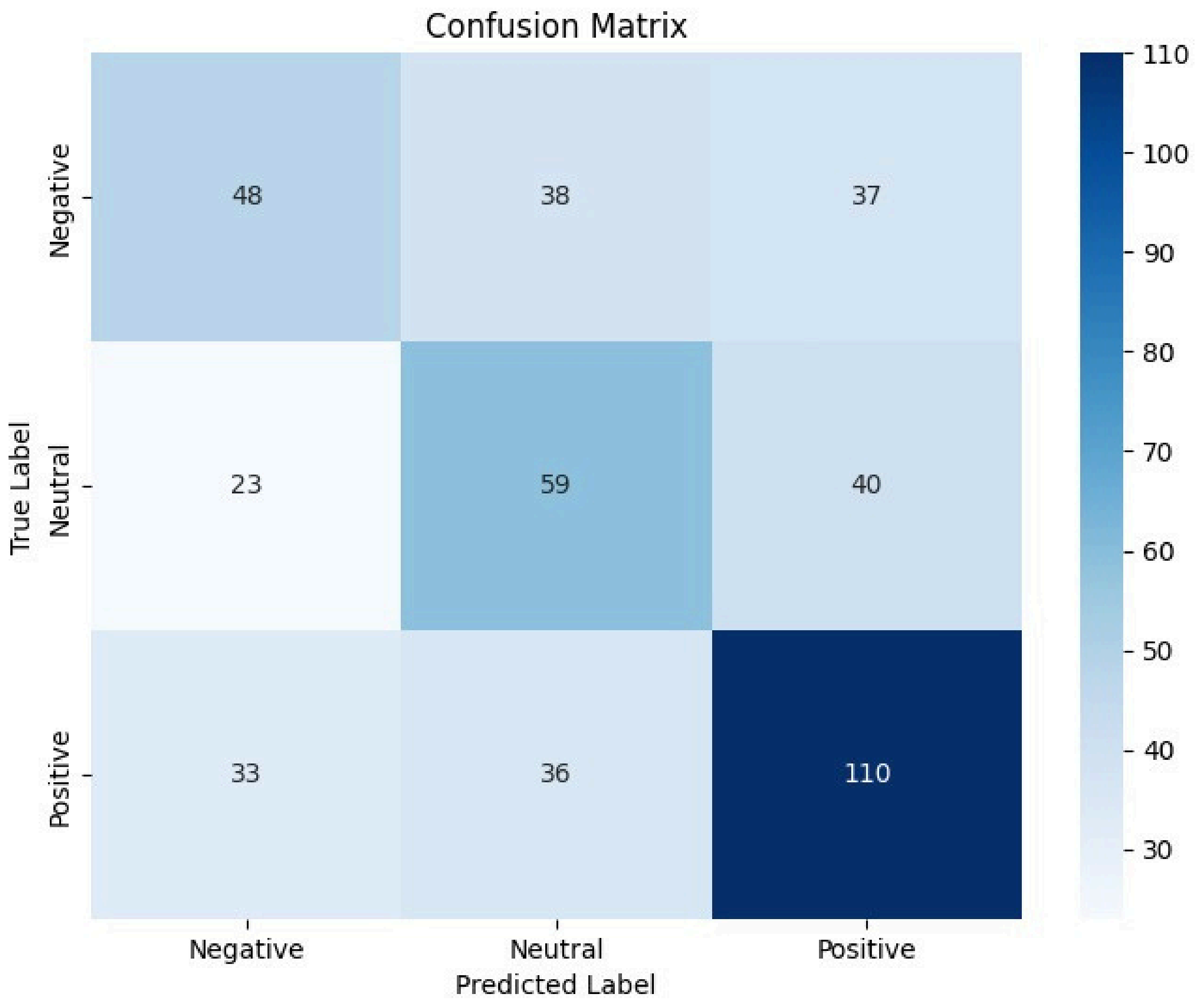


Approach 1

With sentence clusters

Overall Precision: 0.5187690606380431
Overall Recall: 0.5165094339622641
Balanced Accuracy: 0.5058510425
Test Accuracy: 0.516509413719

Since we break the sentences down from documents, there is some context lost in between, thus the sub-optimal results

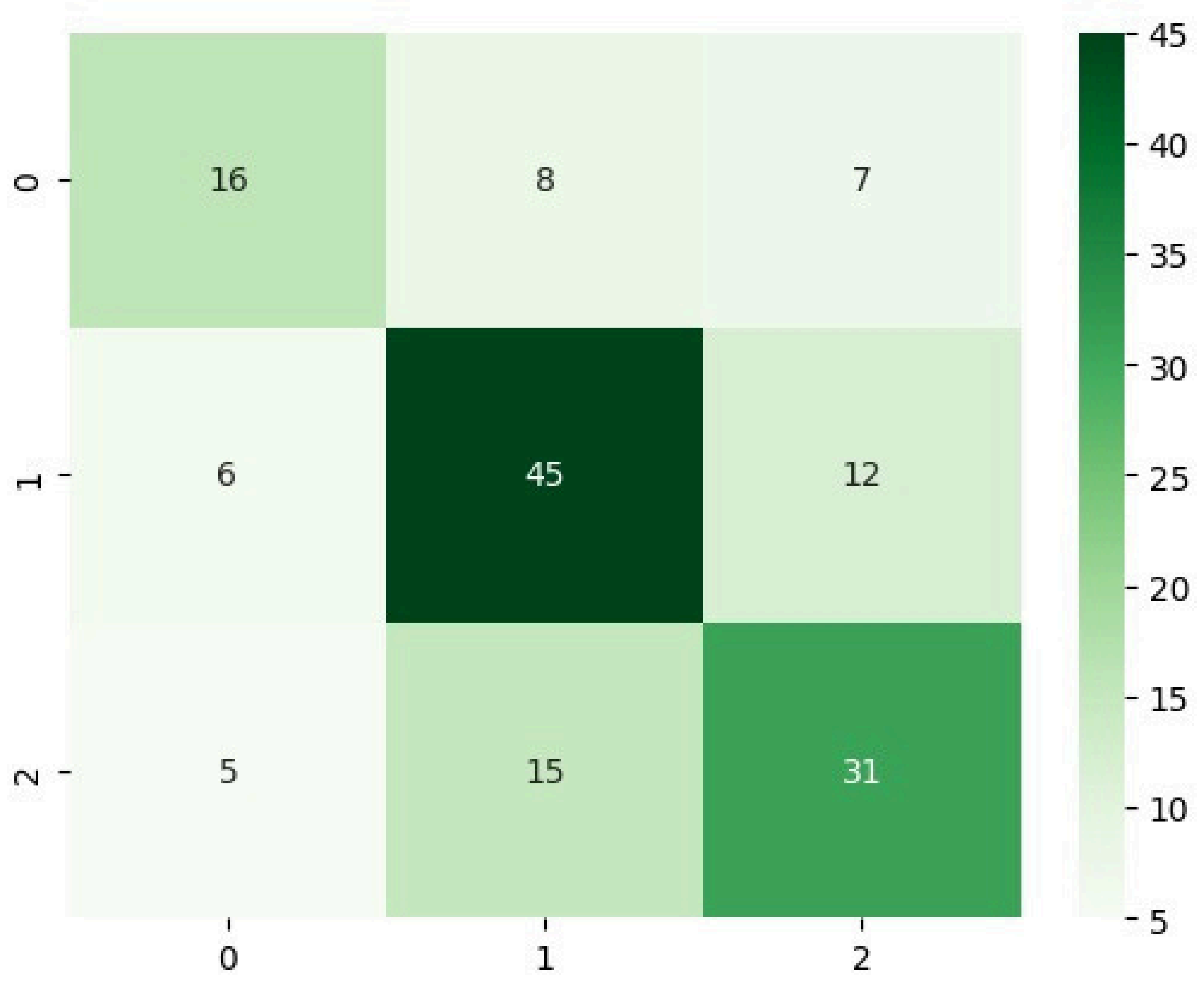


Approach 2

Maintaining the document integreity

Precision: 0.6672748386080931
Recall: 0.6689655172413793
Balanced Accuracy: 0.63920765438792

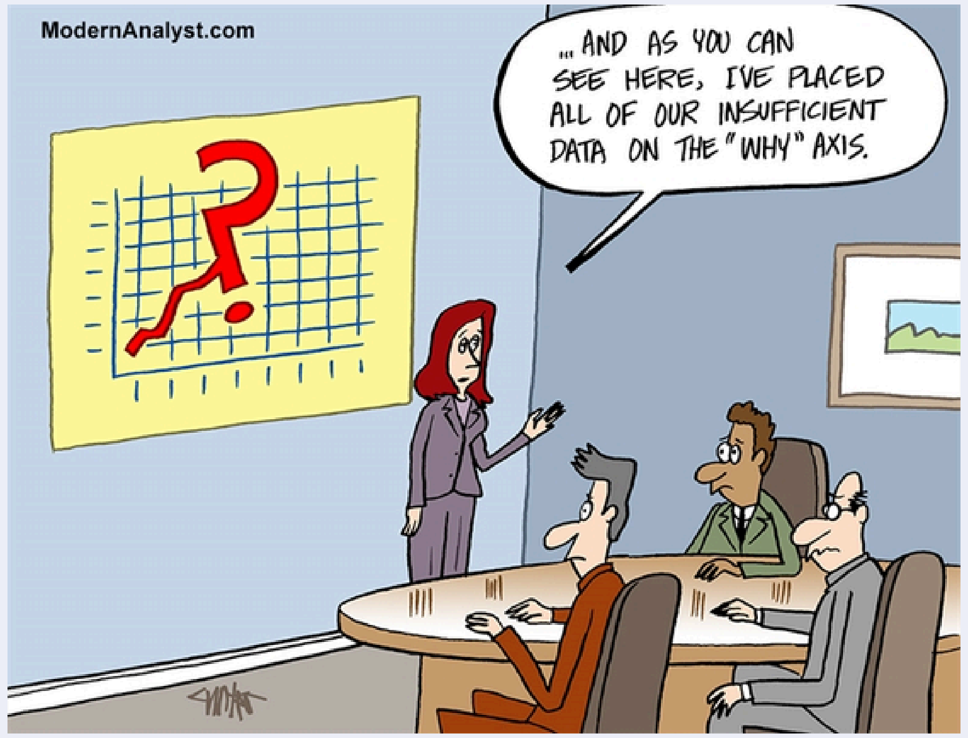
By not breaking the sentences down from documents and preserving the context, we were able to achieve relatively better results shown above



Challenges faced :(

01 Very few data points

We only had MPC data till 1997 with varying frequencies of issues per year.



How did we tackle it?

Data augmentation: Back translation, breaking
Fomc data: The US FED also has had a historical impact on Indian markets. We analysed them as well by adding in our dataset.

Challenges faced :(

02 Lack of existing literature in Indian Context

There was little to no data available in the Indian context for this problem statement, especially in case of a classification problem

How did we tackle it?

Spent time on :
selective feature extraction and data preprocessing,
finalizing the model architecture with no benchmarks.

Challenges faced :(

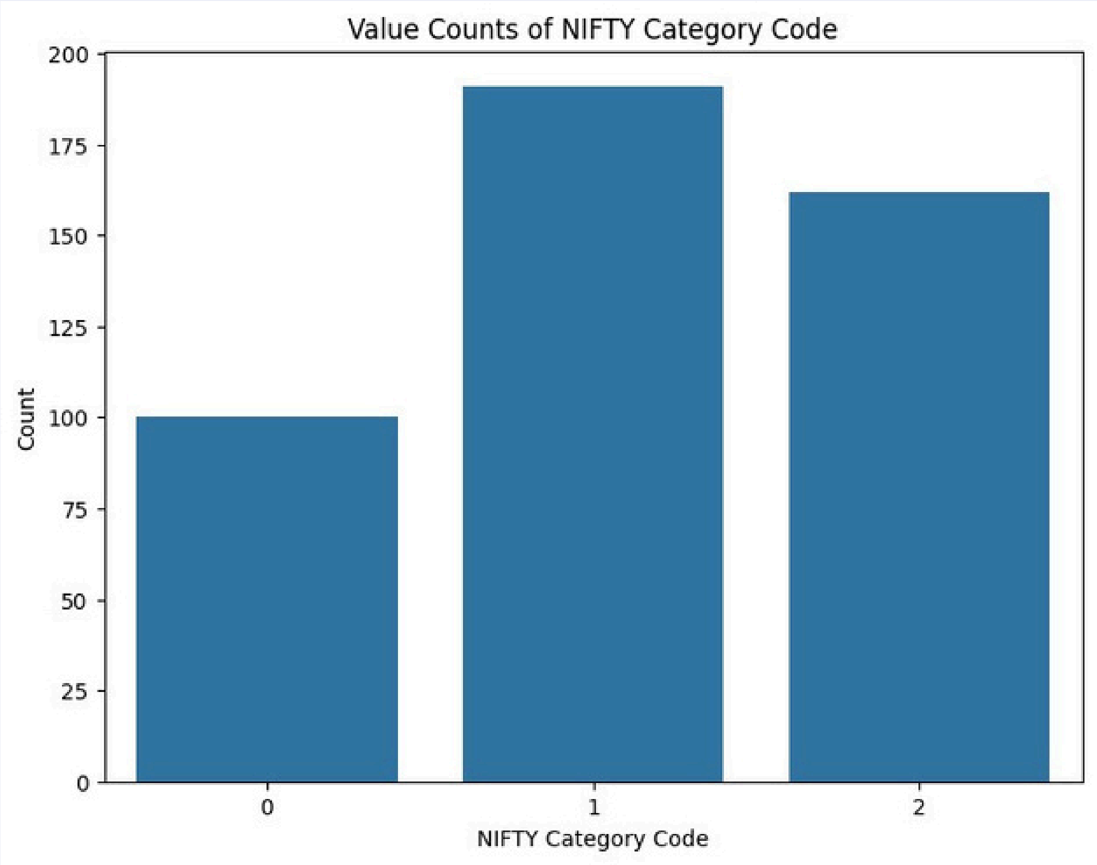
03

Class imbalance

We could not just arbitrarily choose a threshold of values to categorize our target variable into the 3 categories

How we decided to threshold it

Found that a threshold of 1.75% gave us the most balanced classes, after discounting the offset from generalized growth in NIFTY.



No comic, just a plot :\

Real world application:

As a product for Investment firms

With a richer dataset and more diverse financial metrics, this solution has a scope to scale.
It can include Bank of International Settlements (for India) documents, other international banks' documents.

Purpose: Give an edge to these firms to beat competitors and get ahead of the market.

Threat: Security concerns, risk in investments

"The bad news is, our investment lost 10 million dollars. The good news is, it was your money not mine."

That's a wrap!

Thank you