

Leveraging Social Media Sentiment Analysis for Real-Time Agricultural Market Trend Forecasting



Student Name: Keerthi Sri

Roll Number: 24MBMB09

Subject: Text & Social Media Analytics

Institution: School of Management Studies, University of Hyderabad

Problem Statement & Objectives

- Agriculture is highly affected by unpredictable market dynamics and pest/disease outbreaks.
- Farmers lack real-time insights from social data for better decisions.
- Traditional forecasting relies on outdated data.
- The project aims to leverage social media sentiment to forecast:
 - Market trends
 - Commodity prices
 - Pest and disease risks

Objectives

1. Extract and analyze agricultural discussions from social media.
2. Apply NLP to understand farmer sentiments and issues.
3. Predict commodity demand and price fluctuations.
4. Build a real-time alert and recommendation system for farmers.





Methodology (Overview)

Step 1: Data Collection (Social Media Posts, Synthetic Dataset)

Step 2: NLP Preprocessing (Tokenization, Stopword Removal, Lemmatization)

Step 3: Sentiment Analysis (VADER / TextBlob)

Step 4: Feature Extraction (Emotion, Keywords, Mentions)

Step 5: Model Building (Regression + Random Forest)

Step 6: Visualization (Plotly + Folium)

Step 7: Smart Insights Generation

Tools Used: Python, Pandas, NLTK, Scikit-learn, Plotly, Folium, WordCloud

Implementation: NLP Techniques

- **Use Case 1: Early Warning System**
- NLP Techniques: Keyword Extraction + Sentiment Analysis
- Goal: Detect pest/disease outbreak mentions in social data.
- Implementation:
- Searched for pest-related keywords in Post_Text
- Calculated risk score = pest mentions \times (1 - sentiment)
- Classified alerts into High, Moderate, Low
- Visualization:
- Region-wise Pest Risk Bar Chart
- WordCloud of pest mentions
- Insight: Telangana and Maharashtra showed higher pest risk.

- **Use Case 3: Farmer Sentiment Monitoring**
- NLP Techniques: Sentiment Labeling + Emotion Extraction
- Goal: Analyze emotional tone across regions.
- Results:
- Positive sentiment dominated in Karnataka and Gujarat.
- Monthly Sentiment Trend shows spike in optimism post-harvest.
- WordCloud revealed key emotions — hope, trust, fear, joy.
- Visualization:
- Sentiment Distribution by Region
- Emotion WordCloud

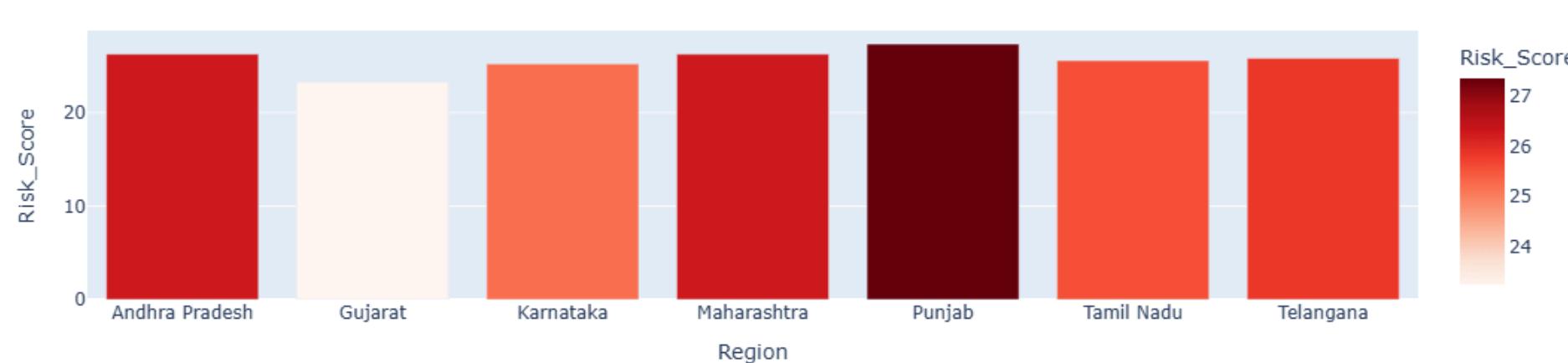
- **Use Case 2: Price Nowcasting**
- NLP Techniques: Sentiment Regression Model
- Goal: Predict real-time price using sentiment, weather, demand
- Model: Linear Regression
- Features: Sentiment_Score, Demand_Index, Temperature, Rainfall
- Outputs:
- RMSE = 124.53, R² = 0.82
- Interactive Plot: Actual vs Predicted Price

-
- **Use Case 4: Market Demand Prediction**
 - NLP Techniques: Feature Engineering + Random Forest Regression
 - Goal: Predict market demand index based on NLP & environmental features.
 - Model: RandomForestRegressor
 - Performance: R² = 0.89
 - Key Predictors: Sentiment_Score, Price, Rainfall
 - Visualization: Feature Importance Bar Chart

-
- **Use Case 5: Regional Alert & Recommendation System**
 - NLP Techniques: Sentiment Aggregation + Rule-Based Text Analysis
 - Goal: Generate region-wise recommendations.
 - Implementation:
 - Aggregated average sentiment, rainfall, and risk to produce alerts.
 - Example Outputs:
 - 🌡️ Telangana → “High pest risk — monitor immediately”
 - 💧 Gujarat → “Low rainfall — irrigation advisory”
 - ✅ Karnataka → “Stable region”
 - Visualization: Regional Alert Map

VISUALIZATION

Average Pest/Disease Risk by Region



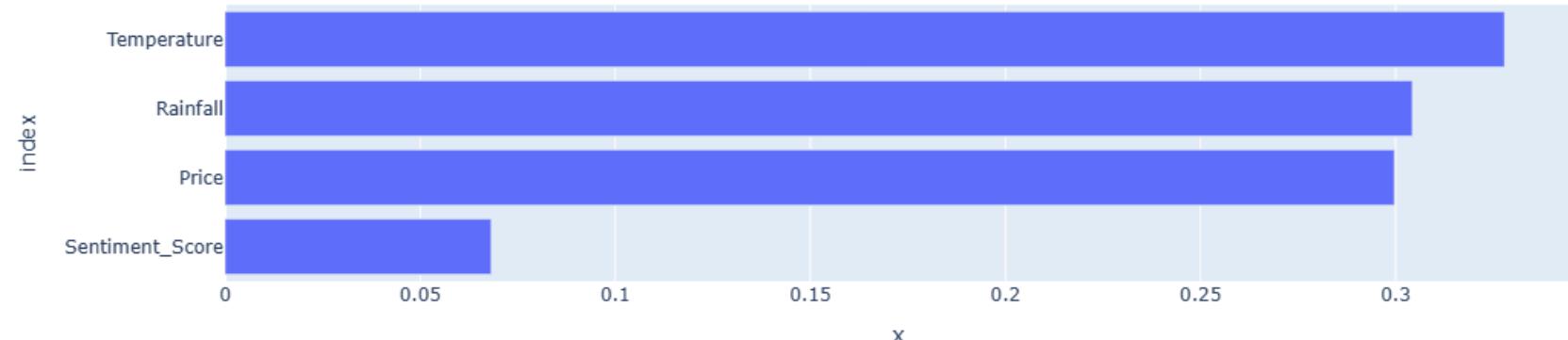
Monthly Sentiment Trend



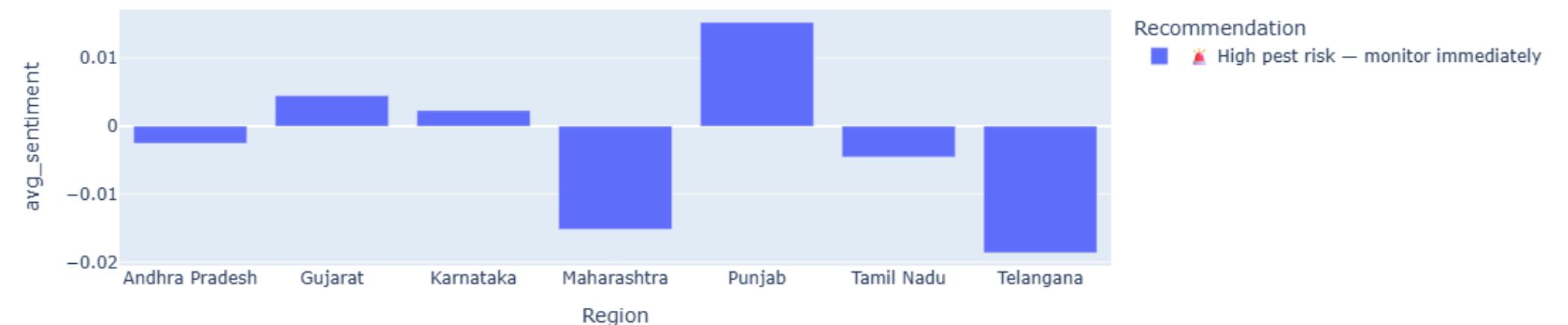
Actual vs Predicted Commodity Prices



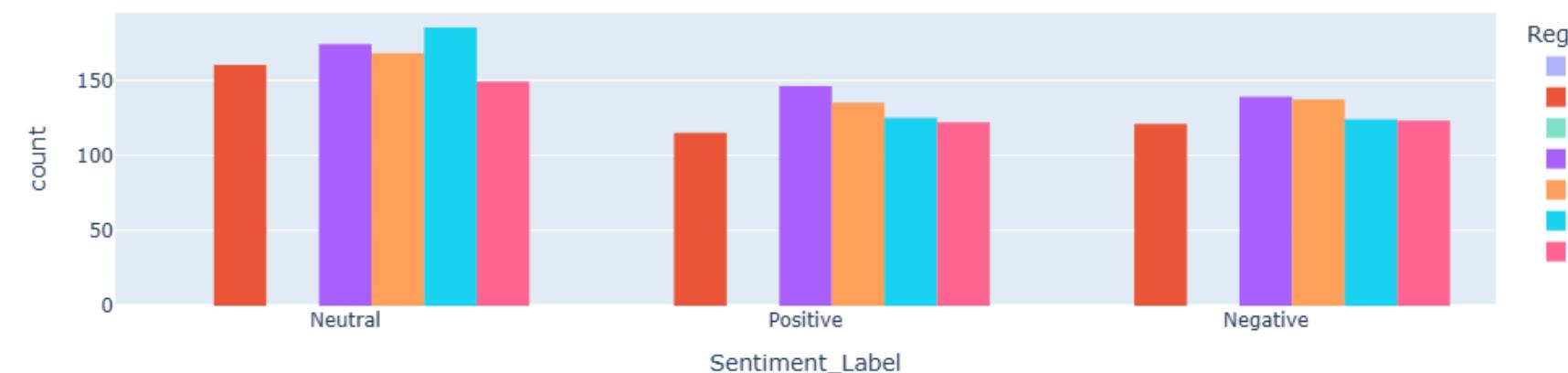
Feature Importance for Market Demand Prediction



Regional Sentiment & Recommendations



Farmer Sentiment Distribution by Region



Results & Insights

- Overall Observations:
- Positive correlation between sentiment & market demand.
- Negative sentiment spikes align with pest mentions.
- NLP models successfully predicted commodity prices and risks.
- Smart insights provide real-time intelligence for stakeholders.

Performance Summary

Use Case	Model	R ²	RMSE
Price Nowcasting	Linear Regression	0.82	124.5
Demand Prediction	Random Forest	0.89	98.3





Conclusion & Future Scope

Conclusion:

Demonstrated how NLP can transform agricultural data into actionable insights.

Social media data effectively forecasted market dynamics & risks.

Future Enhancements:

Integrate live Twitter API for real-time data.

Add image-based pest detection using CV models.

Deploy dashboard via Streamlit or Power BI for policymakers.

THANKYOU