# Employee Sentiment Analysis - Final Report

Author: Kunal Vishwa Sivakumar  
Program: M.S. Applied Data Analytics, Boston University

## 1. Introduction

This report presents the methodology, insights, and outcomes of an Employee Sentiment Analysis project. The goal was to analyze employee email communication using Natural Language Processing (NLP) and machine learning to identify sentiment trends, rank employees by engagement, detect early signs of attrition (flight risk), and build a basic predictive model. The emphasis was on thoughtful interpretation, cross-validation of AI outputs, and the development of explainable insights from real communication data.

## 2. Methodology Overview

The project was implemented in Python using the pandas library for data manipulation, VADER from NLTK for sentiment classification, and scikit-learn for modeling. Sentiment thresholds were based on VADER’s recommended compound score cutoffs (>0.05 = Positive, <-0.05 = Negative, else Neutral), and were validated through manual review of sample emails. Charts and outputs were interpreted critically to ensure meaningful conclusions beyond raw AI-generated results.

## 3. Task Breakdown & Findings

### Task 1: Sentiment Labeling

Approach: VADER sentiment analyzer was applied to email messages using compound score thresholds:  
- Compound > 0.05 → Positive  
- Compound < -0.05 → Negative  
- Otherwise → Neutral  
  
This approach was validated on a subset of emails to ensure it aligned with human judgment in a business context. The results were stored in a new column `sentiment\_label` in test\_labeled.csv.

### Task 2: Exploratory Data Analysis (EDA)

The dataset had no nulls and consisted of sender, timestamp, and message content. Charts were generated to explore trends:  
- Sentiment Distribution: Majority of messages were positive, consistent with a generally satisfied workforce.  
- Monthly Trends: Fluctuations suggested potential event-driven sentiment changes.  
- Top 10 Active Employees: Heavily represented by Enron group members.  
- Word Count Distribution: Most messages were concise, peaking under 100 words.  
  
Each chart was interpreted in context to extract insights, not just visual summaries. All visuals are stored in the `visualization/` folder.

### Task 3: Monthly Sentiment Score

Each message was scored numerically: +1 (Positive), -1 (Negative), 0 (Neutral). These were aggregated by employee and month to generate `monthly\_scores.csv`. This scoring scheme was simple but effective in revealing employee-level sentiment trends over time.

### Task 4: Employee Ranking

Employees were ranked by their total sentiment scores each month. Ties were broken alphabetically.  
- `top\_3\_positive.csv` lists the highest-ranked employees monthly.  
- `top\_3\_negative.csv` lists those with the lowest sentiment scores.  
  
This ranking helped highlight both consistently engaged employees and those showing potential disengagement.

### Task 5: Flight Risk Identification

Employees were flagged as potential flight risks if they sent 4 or more negative messages within any 30-day rolling window. This approach avoids rigid calendar boundaries and captures sustained negative behavior over time. Manual inspection of flagged emails confirmed many expressed stress, burnout, or dissatisfaction. Results were saved in `flight\_risks.csv`.

### Task 6: Predictive Modeling

A Linear Regression model was built to predict an employee’s monthly sentiment score based on:  
- Number of messages sent  
- Average message length  
  
These features were selected for their logical connection to engagement. The model was evaluated using MSE and R²:  
- R² indicated variance explained  
- MSE highlighted potential high-error predictions due to outliers  
  
Results showed moderate predictive power, indicating that additional features (e.g., topic analysis, emotion keywords) would be beneficial.

## 4. Key Insights

- Positive sentiment was dominant, indicating overall satisfaction.  
- Some employees consistently ranked highly or poorly, flagging stable behavior trends.  
- Flight risks often had extended negative streaks, validating the rolling window method.  
- The model captured base-level patterns, but richer inputs are needed for robust forecasting.

## 5. Recommendations

- Proactively engage with employees flagged as potential flight risks.  
- Expand features with topic/emotion tagging for future predictive modeling.  
- Implement real-time dashboards to track sentiment dynamics over time.  
- Cross-validate AI outputs regularly to prevent misclassification bias.

## 6. Deliverables Summary

Notebook: Employee.ipynb  
Outputs: CSV files under outputs/  
Visuals: Plots under visualization/  
Readme: README.md with complete summary