Sentiment Analysis Assignment – Summary Document

1. Overall Strategy

My strategy was to take the raw call transcripts and separate **agent** and **borrower** utterances. For each call, I concatenated the utterances by speaker and computed an overall sentiment score per speaker. These regression-like sentiment scores (in the range [-1,1]) were then compared with the **categorical call disposition** (positive/negative). The aim was to see whether simple sentiment features could explain or predict call outcomes.

---

2. Hypotheses

- Calls with **negative borrower sentiment** (e.g., frustration, refusal) would correlate with **negative dispositions** such as "No Pay".
- Calls where the **agent maintained a positive sentiment** might mitigate negative borrower sentiment and improve outcomes.
- Borrower sentiment would be a stronger predictor of call outcomes than agent sentiment.

---

3. Evaluation Method & Rationale

I evaluated the sentiment features with:

- **Descriptive statistics**: comparing mean sentiment across positive vs negative outcomes.
- **Correlation analysis**: point-biserial correlation between sentiment scores and binary outcome.
- **Predictive modeling**: logistic regression using agent and borrower sentiment as features.
- **Classification metrics**: accuracy, ROC AUC, F1, and feature importance from the model.

`DecisionTree` was chosen because it is a simple, interpretable baseline that shows how much signal exists in the sentiment scores.

---

4. Findings

- **Borrower sentiment** was far more predictive than agent sentiment.
    - `DecisionTree` feature importances:
        - Agent sentiment: **0.15**
        - Borrower sentiment: **0.85**
- **Model performance (n=2000 calls):**
    - Accuracy: **0.6225**
    - ROC AUC: **0.6653**
    - F1 Score: **0.627**

Classification report:

```
              precision    recall  f1-score   support

           0       0.63      0.61      0.62      1000
           1       0.62      0.63      0.63      1000

    accuracy                           0.62      2000
   macro avg       0.62      0.62      0.62      2000
```

Correlation results (`pointbiserialr`):

- Agent sentiment correlation with outcome: r = 0.042, p = 0.0576** (not significant).
- Borrower sentiment correlation with outcome: r = 0.254, p ≈ 9.66e-31** (strong and significant).

These results indicate that borrower sentiment contains a predictive signal, whereas agent sentiment remains relatively flat across outcomes.

`The` **`best working solution`** `for this dataset was a simple TextBlob polarity`

---

5. Extensions with More Time

With more time, I would:

- Use **transformer-based sentiment models** (e.g., FinBERT, RoBERTa) tuned for financial/conversational text to capture more nuance.
- Analyze **sentiment trajectories** (how borrower sentiment changes over the call).
- Add **interaction features**, such as sentiment gap (agent–borrower).
- Explore **topic modeling** or **intent recognition** in conjunction with sentiment analysis.
- Evaluate more advanced classifiers (Random Forest, XGBoost) for nonlinear patterns.

---

6. Next Steps Before Production

Before deploying:

- Validate on a **larger, more diverse dataset**.
- Compare against **human-labeled sentiment annotations**.
- Ensure **robust preprocessing** for anonymized text and speech artifacts.
- Test **latency and scalability** for running across thousands of daily calls.
- Implement **human-in-the-loop monitoring** to improve and calibrate sentiment predictions over time.