Sentiment Analysis Results

# Sentence Sentiment Analysis

## Reports

### Classification Report (Top Table)

**This shows precision, recall, F1-score, and support for each class.**

* **Negative (61 samples)**
  + **Precision = 0.98 → When the model predicts *negative*, 98% are correct.**
  + **Recall = 0.92 → Of all true *negative* cases, 92% are captured (8% missed, misclassified as *positive*).**
  + **F1-score = 0.95 → Balance between precision & recall.**
* **Neutral (278 samples)**
  + **Precision = 1.00 → Every time the model predicts *neutral*, it’s always correct.**
  + **Recall = 0.99 → Almost all true *neutral* cases are identified (only 1% missed).**
  + **F1-score = 0.99 → Very strong performance here.**
* **Positive (114 samples)**
  + **Precision = 0.93 → When predicting *positive*, 93% are correct (7% false positives).**
  + **Recall = 0.98 → Almost all true *positives* are captured.**
  + **F1-score = 0.95 → Solid balance.**

**Overall metrics**

* **Accuracy = 0.98 → Out of all 453 test samples, 98% are correct.**
* **Macro avg = averages metrics equally across classes → F1 = 0.96**
* **Weighted avg = averages weighted by class size → F1 = 0.98 (higher because *neutral* dominates with 278 samples and is predicted very well).**

### Confusion Matrix (Bottom Plot)

**This shows actual vs predicted counts:**

* **Row = True class, Column = Predicted class**
  + **Negative (61 true):**
    - **56 correctly predicted as negative**
    - **5 misclassified as positive**
  + **Neutral (278 true):**
    - **274 correctly predicted**
    - **4 misclassified as positive**
  + **Positive (114 true):**
    - **112 correctly predicted**
    - **1 misclassified as negative**
    - **1 misclassified as neutral**

**👉 Errors are very few and mostly between negative ↔ positive or neutral ↔ positive (natural since they’re semantically closer).**

### Key Takeaways

1. **The model performs exceptionally well (98% accuracy, F1 > 0.95 across all classes).**
2. **Neutral is almost perfectly predicted (possibly because it’s the majority class).**
3. **Negative has slightly lower recall (0.92) → some negatives get mistaken as positives.**
4. **Overall balance is strong, meaning class imbalance is not severely hurting performance.**

## Business/Real-World Interpretation

### Why These Metrics Matter in Business Context

**1. Negative Class (Customer complaints, risk alerts, bad financial news)**

* **Recall = 0.92** → The model catches **92% of negatives**, but misses 8%.
* **Business impact:** Missing a negative (false negative) could mean overlooking a **serious issue** (e.g., financial risk, compliance violation, or customer dissatisfaction).
* In risk-sensitive industries (finance, insurance, healthcare), **high recall is crucial** → you’d rather flag too many negatives than miss one.

**2. Neutral Class (Normal or uninformative content)**

* **Precision = 1.00, Recall = 0.99** → The model is almost perfect here.
* **Business impact:** Neutral predictions are low risk. If the model mistakes a few, it doesn’t matter much since these don’t trigger critical business actions.
* High accuracy here is good, but **less critical** compared to negative or positive.

**3. Positive Class (Opportunities, good sentiment, positive financial signals)**

* **Precision = 0.93, Recall = 0.98** → Very strong performance.
* **Business impact:** If this model is used for **market sentiment or customer engagement**, then missing positives (false negatives) could mean **overlooking opportunities** (e.g., happy customers you could upsell, bullish market news).
* Good recall ensures you don’t miss many positives, but slightly lower precision means a few false positives (things flagged as positive but not truly so).

### Metric Priorities Depend on Use Case

* **If used for risk detection (compliance, fraud, complaints):**
  + **Maximize recall for negatives** → better to catch every possible negative, even if some are false alarms.
  + A false positive is annoying, but a false negative could cost millions.
* **If used for opportunity mining (marketing, investments, client engagement):**
  + **Maximize recall for positives** → don’t miss good opportunities.
  + Slightly lower precision is acceptable, because reviewing extra "positives" is less costly than missing a big chance.
* **If used for overall sentiment monitoring (balanced insights):**
  + The current model is already well-balanced → 98% accuracy, F1 ~0.95+ across classes means it’s strong enough for dashboards, analytics, and trend reporting.

### Summary in Business Terms:

* The model is **excellent overall**, but depending on your business goal, you might want to **tune the loss function or decision thresholds** to prioritize recall for *negatives* (risk detection) or *positives* (opportunity spotting).
* For general analytics, this model is already production-ready.