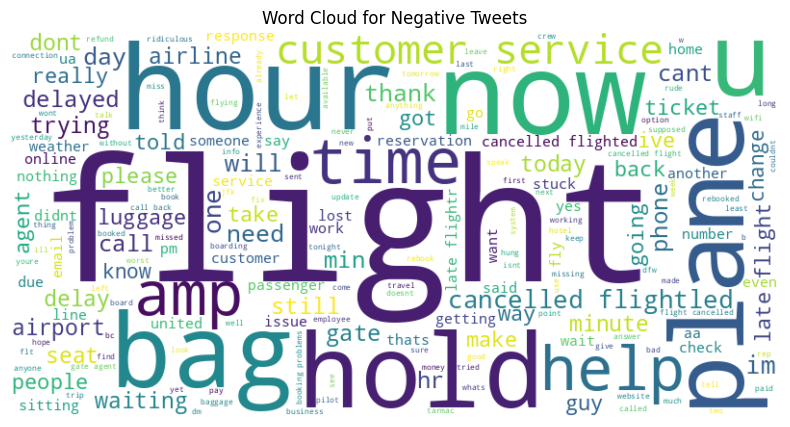
**Lab 10**

**Objective**

The goal of this task was to build a sentiment classification model for airline tweets. The model predicts whether a tweet's sentiment is positive, neutral, or negative.

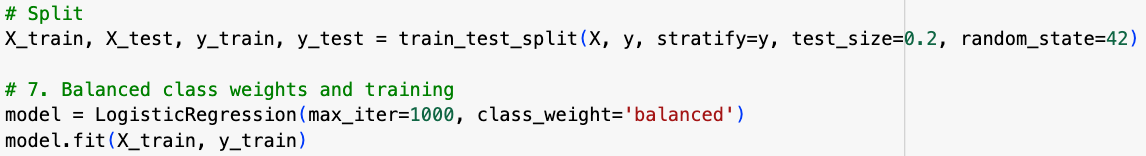
**Word Cloud (Negative Tweets)**

The following word cloud visualizes the most frequent words found in **negative sentiment** tweets.



**Model Summary**

* **Model Used:** Logistic Regression
* **Feature Extraction:** TF-IDF with bi-grams (max\_features=5000)
* **Class Imbalance Handling:** class\_weight='balanced'
* **Train/Test Split:** 80/20 (stratified)



**Classification Report**

precision recall f1-score support

negative 0.89 0.77 0.83 1835

neutral 0.51 0.72 0.60 620

positive 0.68 0.68 0.68 473

accuracy 0.74 2928

macro avg 0.69 0.72 0.70 2928

weighted avg 0.78 0.74 0.75 2928

**Final Accuracy: ~74.38%**

**Sample Predictions**

|  |  |
| --- | --- |
| **Tweet** | **Predicted Sentiment** |
| I'm so happy with the service from United! | positive |
| This was the worst flight experience ever | negative |
| The plane arrived late but the staff was okay. | negative |
| Good food, staff were terrible. | negative |
| Bad food, staff were good. | positive |

=> The model tends to weigh **negative cues** more heavily in compound sentences (e.g., "staff were terrible" drives the classification even if "good food" is present).

**Conclusion**

This model performs reasonably well in identifying airline sentiment, especially for negative reviews, which dominate the dataset. The TF-IDF + Logistic Regression pipeline is efficient and interpretable. However, class imbalance and subtle phrasing still challenge neutral/positive predictions.