**Nykaa Product Reviews - Sentiment Analysis Report**

**1. Project Overview**

This project analyses Nykaa product reviews to extract sentiment insights using traditional machine learning and **BERT-based deep learning models**. An Aspect-Based Sentiment Analysis (ASBA) was also conducted to understand customer sentiments towards different product aspects.

**2. Dataset Preprocessing**

* The dataset contains **61,284 reviews** with product details, ratings, and textual reviews.
* Missing values were handled, and text data was cleaned using **regular expressions**.
* **Sentiment Labels** were assigned based on ratings:
  + **Positive** (4+ stars)
  + **Neutral** (3 stars)
  + **Negative** (<3 stars)
* **TF-IDF Vectorization** was used to convert text data into numerical form for machine learning models.

**3. Traditional Machine Learning Models - Performance**

|  |  |  |
| --- | --- | --- |
| **Model** | **Accuracy** | **F1-Score** |
| **Logistic Regression** | 74.6% | 74.5% |
| **SVM** | 74.0% | 74.0% |
| **Random Forest** | 69.2% | 69.0% |
| **XGBoost** | 69.7% | 68.5% |

🔹 **Key Observations:**

* Logistic Regression and SVM performed best among traditional models.
* Random Forest and XGBoost struggled, likely due to lack of contextual understanding.

**4. BERT-Based Sentiment Analysis - Performance**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metric** | **Negative** | **Neutral** | **Positive** | **Overall Accuracy** |
| **Precision** | 83% | 56% | 86% | **77.7%** |
| **Recall** | 77% | 56% | 90% | **77.7%** |
| **F1-Score** | 80% | 56% | 88% | **77.8%** |

🔹 **Key Observations:**

* BERT outperformed all traditional models, achieving **higher accuracy (77.7%)**.
* BERT captured **sentiment-heavy words** like "amazing," "not worth," and "disappointing," leading to better predictions.

**5. Attention Mechanism Insights**

* **BERT’s Attention Heatmaps** showed that sentiment-heavy words received higher attention.
* **Neutral sentiments were harder to classify**, as they had weak focus points.
* **Key learning:** Traditional models fail to understand **word relationships**, while BERT excels in **contextual understanding**.

**6. Aspect-Based Sentiment Analysis (ASBA) Insights**

**Top Aspects Discussed by Customers**

|  |  |
| --- | --- |
| **Aspect** | **Mentions** |
| **Price** | 1,650+ |
| **Quality** | 1,600+ |
| **Fragrance** | 850+ |
| **Packaging** | 450+ |

🔹 **Key Observations:**

* **Price and Quality are most discussed**, suggesting these are critical factors influencing purchase decisions.
* **Fragrance and Packaging are secondary concerns** but still impact customer satisfaction.
* **Business Impact:** Brands should focus on **improving quality-to-price ratio** to increase positive sentiment.

**7. Conclusion & Business Recommendations**

**BERT outperformed traditional ML models** and is better suited for sentiment analysis due to its deep contextual understanding.

**Price and Quality drive customer opinions**, making them crucial for product positioning. **Aspect-Based Sentiment Analysis (ASBA) helps identify customer concerns**, allowing brands to improve key features.