🎤 Presentation Speech (8 Minutes)

---

1. Introduction (1 minute)

Good [morning/afternoon], everyone.
Our project is about **Consumer Review Analysis on Clothing Products**.

We worked with a dataset from Kaggle that contains:

- customer review text,
- ratings,
- clothing categories,
- material and durability details.

Our main goal was to predict customer ratings using **machine learning models**.
We focused on **Support Vector Machines (SVM)** and **K-Nearest Neighbours (KNN)** to find patterns in reviews and understand customer satisfaction.

---

2. Literature Survey (1 minute)

Online reviews are very important today because customers use them before buying, and companies use them to improve products.

Since reviews are so many, **machine learning** is often used to analyze them.

From past research:

- **SVM** works well with high-dimensional data like text.
- **KNN** is simple and good for finding similarities.

But most studies only use review text. Our project adds **extra product features** like category and durability, which improves accuracy.

---

3. Methodology (1.5 minutes)

The steps we followed were:

1. **Data Preprocessing**
   - Cleaned the text (removed stopwords, symbols).
   - Converted text into numbers using **TF-IDF vectorization**.
   - Encoded categorical product features.
2. **Data Splitting**
   - Divided into training and test sets.
3. **Model Training**
   - Trained **SVM** with different kernels.
   - Trained **KNN** with different neighbor values.
4. **Evaluation**
   - Compared results using accuracy, precision, recall, and F1-score.

---

4. Implementation (1 minute)

We implemented the project in **Python using Jupyter Notebook**.

The libraries used were:

- **Pandas, NumPy** for data,
- **Scikit-learn** for models,
- **Matplotlib/Seaborn** for graphs.

Both models were trained on the same dataset so we could compare them fairly.

---

5. Results & Discussion (2 minutes)

Now, the results.

- **SVM performed better** than KNN.
- The accuracy and F1-score of SVM were higher.
- This is because SVM is stronger with high-dimensional text data.

KNN still worked, but it was less accurate when the dataset grew larger.

Key insight: Adding product features along with text helped improve prediction quality.

For businesses, this means they can quickly analyze thousands of reviews and understand what customers like or dislike.

6. Conclusion (1 minute)

To conclude:

- We applied **SVM and KNN** to analyze clothing reviews.
- **SVM gave the best results**.
- Using both review text and product details improved accuracy.

In the future, we can try:

- Deep learning methods like LSTM or Transformers,
- Multilingual review analysis,
- Or larger datasets.

This project shows the value of machine learning in **understanding customer opinions** and helping fashion retailers improve products.

---

7. Closing (30 seconds)

We would like to thank our lecturers for guidance and Kaggle for the dataset.

Thank you for your attention. We are happy to answer any questions.