Streamlining Product Categorization for E-commerce Using LoRA

Sai Prerana Mandalika, Pranav Sudhir Patil

1. Introduction

Product categorization in e-commerce is a crucial component for enhancing user experience and operational efficiency. Accurate categorization helps customers navigate vast product catalogs, improves search functionality, and increases sales through better discoverability. However, manual categorization faces challenges such as inconsistency, high costs, and impracticality at scale. To address these issues, this project focuses on fine-tuning pre-trained Large Language Models (LLMs) using Low-Rank Adaptation (LoRA), a cost-effective and computationally efficient technique.

By leveraging LoRA, we aim to balance the trade-offs between computational efficiency and model performance. This approach reduces the number of trainable parameters, enabling fine-tuning on resource-constrained devices without compromising accuracy. Our project evaluates LoRA-enhanced models against traditional fine-tuning techniques for automated product categorization. Key findings demonstrate significant improvements in scalability and resource utilization, laying the groundwork for practical implementations in the e-commerce industry.

---

2. Problem Definition and Algorithm

2.1 Task Definition

The problem addressed is the automated categorization of products in e-commerce platforms. Given a dataset of product descriptions and their respective categories, the task involves classifying each product into its correct category.

Inputs:

- Textual product descriptions.

- Metadata such as product titles.

Outputs:

- Predicted category labels for each product.

This problem is significant as it directly impacts user experience, inventory management, and targeted marketing strategies. Automating this process ensures consistency, scalability, and reduced operational costs.

2.2 Algorithm Definition

To tackle this problem, we utilized the LoRA framework for fine-tuning pre-trained LLMs, including GPT-2, BERT, and DistilRoBERTa. LoRA modifies a subset of model parameters by introducing low-rank matrices, significantly reducing computational overhead while preserving pre-trained knowledge.

Algorithm Steps:

1. **Dataset Preparation**:
    - Source: Amazon Reviews Dataset 2023.
    - Categories include "Fashion," "Grocery," "Handmade," and others.
    - Preprocessing: Tokenization, label encoding using Hugging Face's LabelEncoder class.
2. **Model Initialization**:
    - Load pre-trained models (i.e., GPT-2, BERT, and DistilRoBERTa).
    - Define LoRA-specific parameters (e.g., rank, alpha, dropout).
3. **Training**:
    - Split the dataset into training, validation, and testing sets.
    - Fine-tune the model with LoRA configurations.
    - Hyperparameter tuning (learning rate, batch size, weight decay).
4. **Evaluation**:
    - Evaluate performance using metrics: accuracy, precision, recall, and F1-score.

3. Experimental Evaluation

3.1 Methodology

We evaluated the effectiveness of LoRA against traditional fine-tuning based on the following hypotheses:

- LoRA achieves comparable categorization accuracy to full fine-tuning.
- LoRA significantly reduces training time and memory requirements.

Evaluation Criteria:

- **Dependent Variables**: Accuracy, precision, recall, F1-score.
- **Independent Variables**: Model architecture, LoRA parameters, dataset splits.

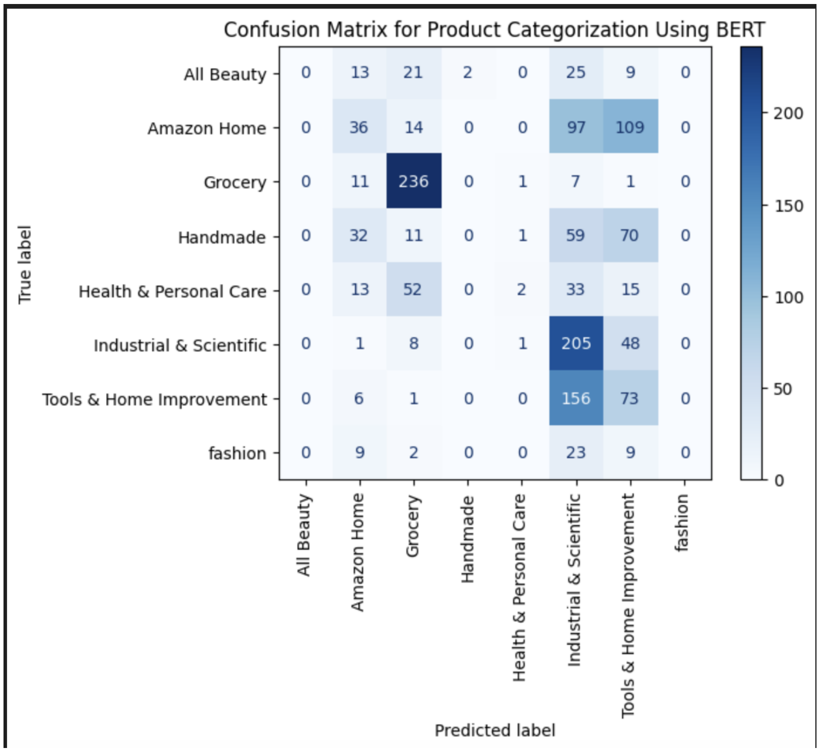
Data:

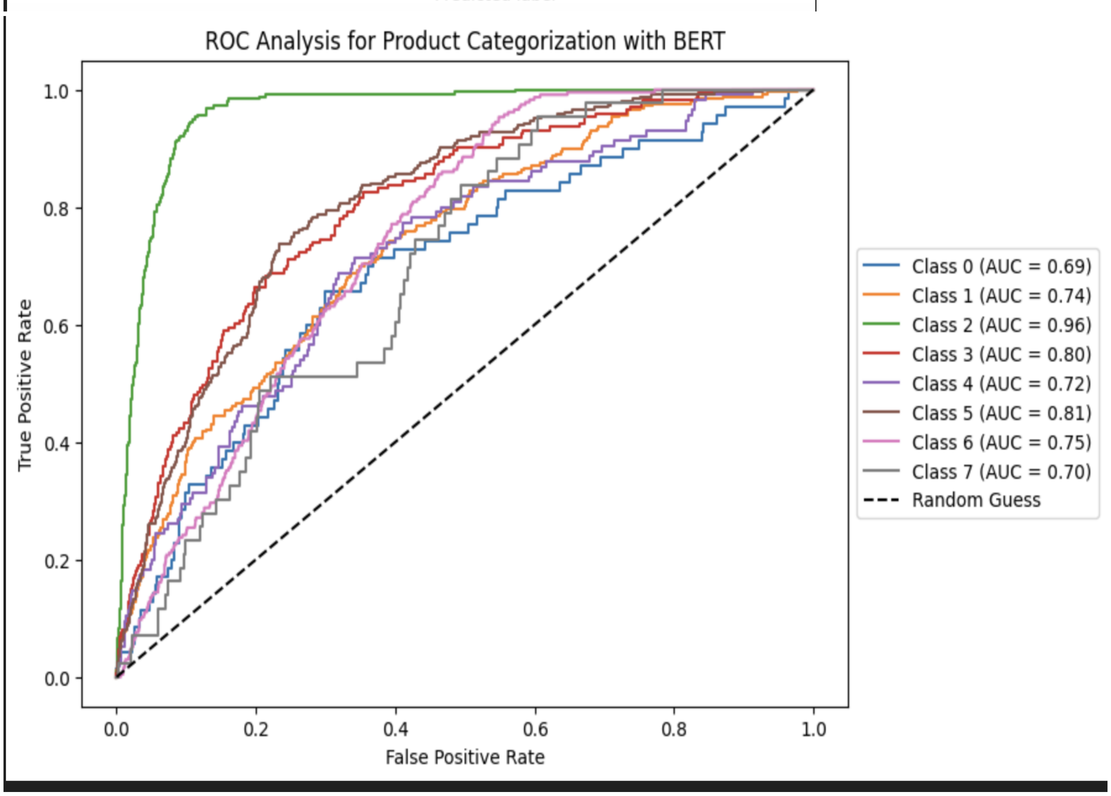
- Training Set: 80% of the dataset.
- Validation Set: 10% of the dataset.
- Testing Set: 10% of the dataset.

Comparative analyses were performed against full fine-tuning techniques to assess computational efficiency and categorization accuracy.

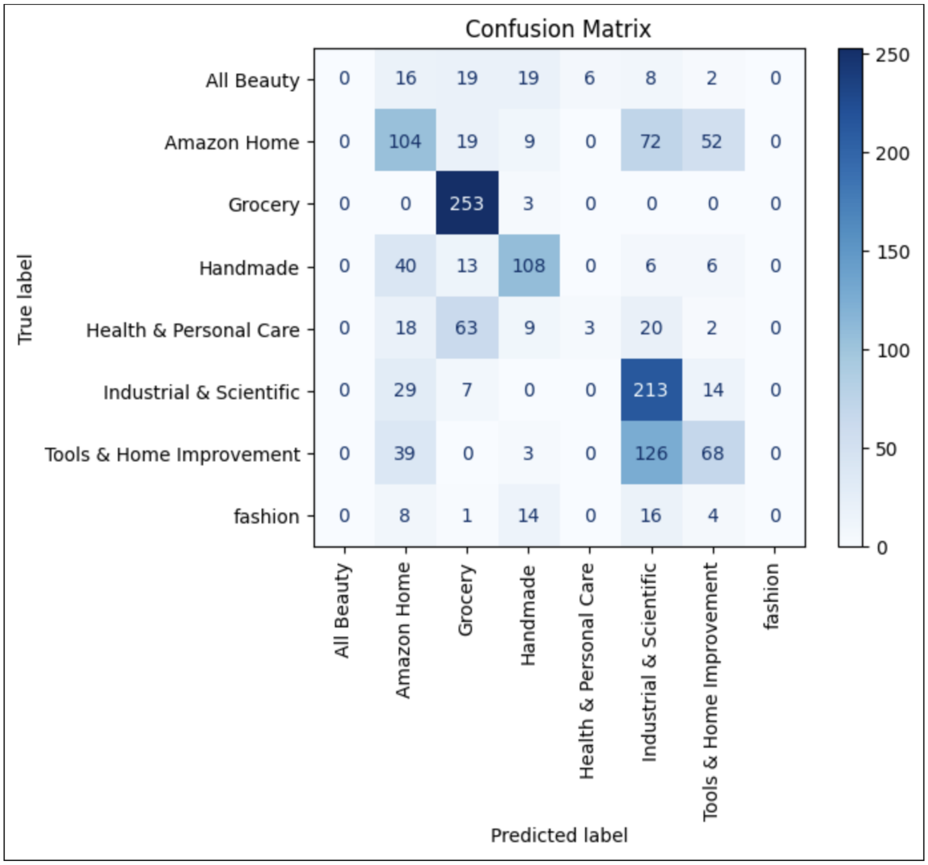
3.2 Results

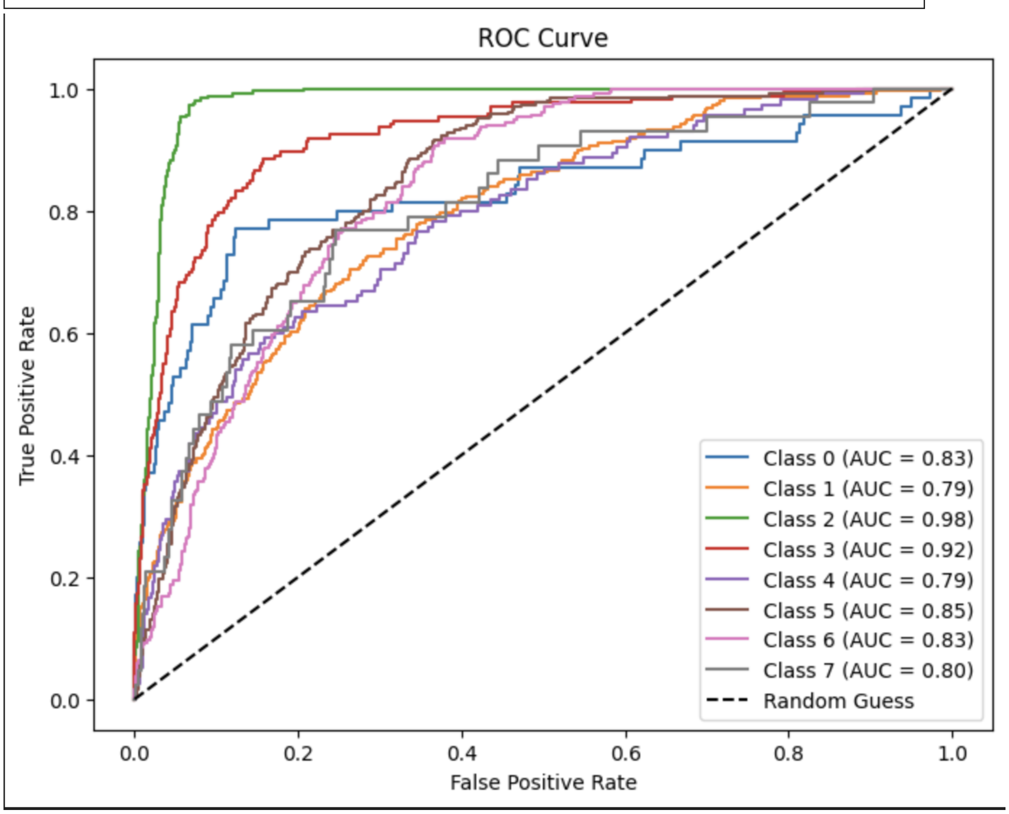
Plot Metrics of Bert Model



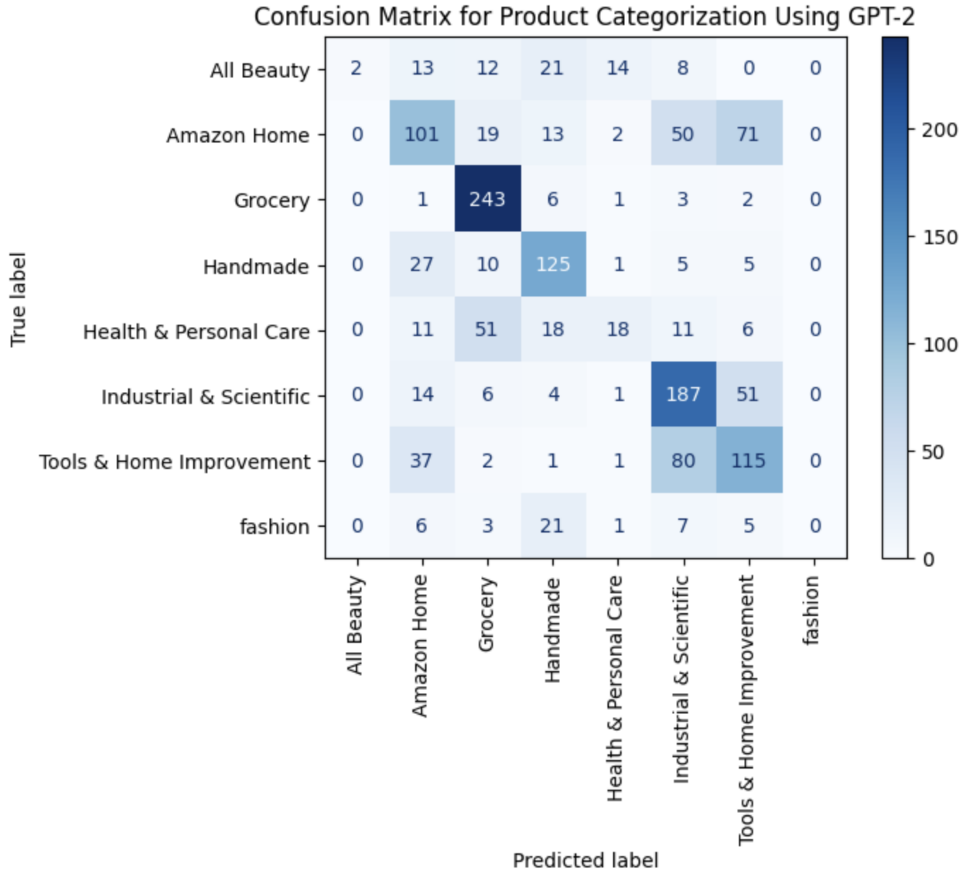


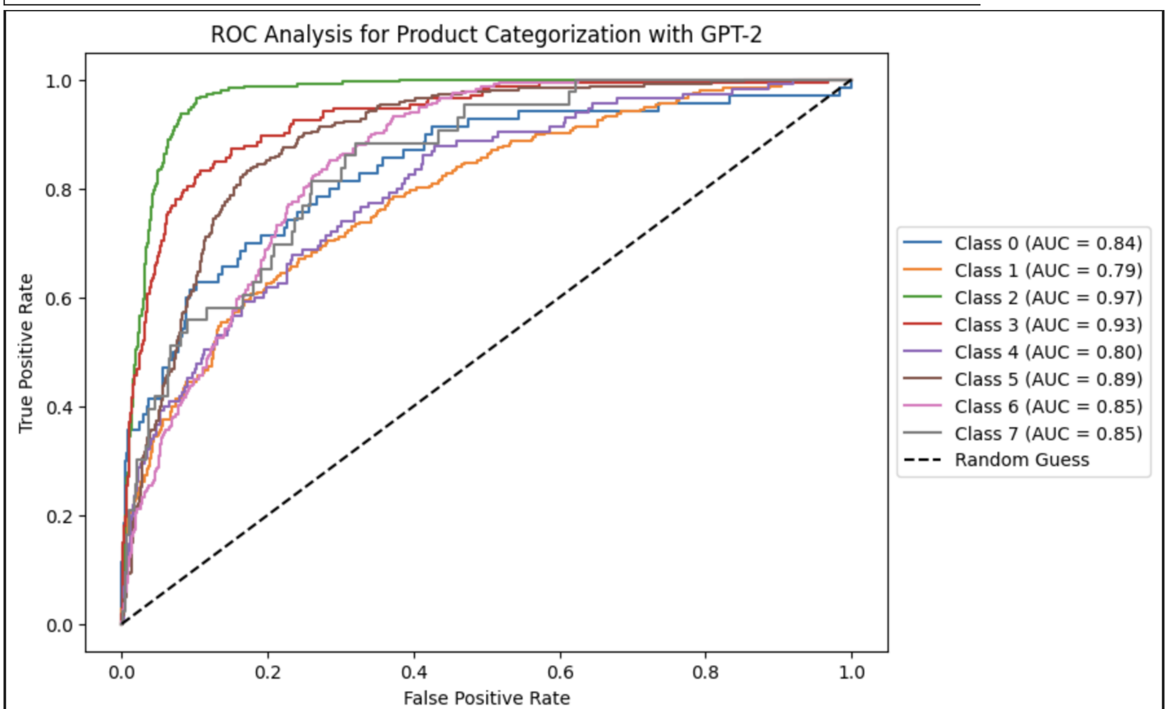
Plot Metrics of Distil-Roberta





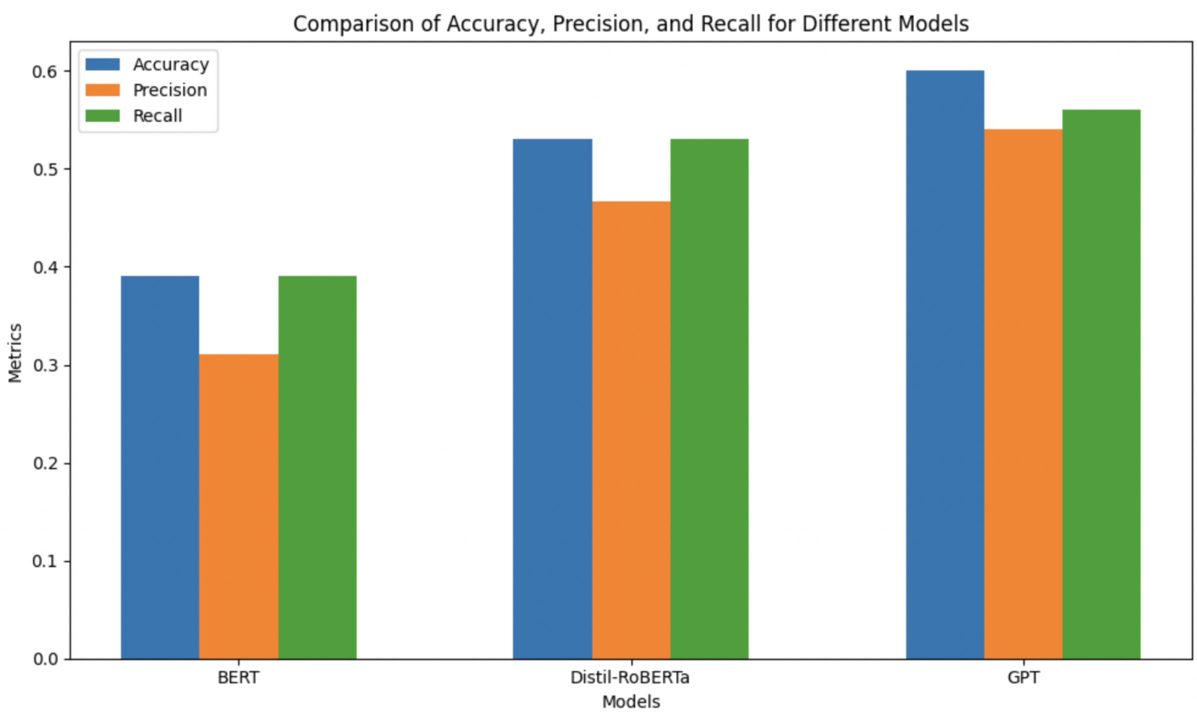
Plot metrics of GPT-2





GPT-2 outperforms BERT in most aspects of the product categorization task. The best-performing class for both models is **Grocery (Class 2)**, with GPT-2 slightly improving the AUC from 0.96 to 0.97. While BERT's worst-performing class is **All Beauty (Class 0)** with an AUC of 0.69, GPT-2 improves this to 0.84, and its lowest AUC is for **Amazon Home (Class 1)** at 0.79, showing an overall boost in performance. GPT-2 also demonstrates higher diagonal dominance in the confusion matrix, indicating fewer misclassifications compared to BERT. Significant improvements are seen in specific classes, such as **Handmade (Class 3)**, where AUC increases from 0.80 to 0.93, and **Industrial & Scientific (Class 5)**, where AUC rises from 0.81 to 0.89. However, both models continue to struggle with overlapping categories, such as **Amazon Home → Handmade** and **Tools & Home Improvement → Industrial & Scientific**, though GPT-2 reduces the frequency of such errors. Overall, GPT-2 achieves better AUC scores and classification accuracy, making it the stronger model for this task, though further fine-tuning and feature engineering are needed for challenging and semantically similar categories.

This plot is a **comparative analysis of Accuracy, Precision, and Recall** metrics for three different models: **BERT**, **Distil-RoBERTa**, and **GPT-2**. GPT-2 consistently outperforms the other two models across all three metrics, making it the most effective model in this comparison.



3.3 Discussion

The results support our hypothesis that LoRA enables efficient and accurate product categorization. GPT-2 emerged as the best-performing model, demonstrating LoRA's ability to enhance computational efficiency without sacrificing performance. Limitations include lower precision for less-represented categories, highlighting the need for handling class imbalance.

4. Related Work

Previous studies have explored fine-tuning LLMs for text classification. For instance:

- Traditional fine-tuning techniques achieve high accuracy but are resource-intensive.
- LoRA, as proposed by Hu et al., introduces a novel approach for efficient parameter tuning.

Our work differentiates by applying LoRA specifically to e-commerce product categorization, addressing unique challenges such as class imbalance and scalability.

---

5. Future Work

Key areas for improvement include:

- Exploring additional LLMs such as RoBERTa, T5, and ALBERT for enhanced performance.
- Incorporating advanced feature representation methods for better semantic understanding.
- Implementing cross-validation for robust hyperparameter tuning.
- Implementing stratified sampling to make sure equal class distribution between train and test dataset.

---

6. Conclusion

This project demonstrated the potential of LoRA-enhanced LLMs for efficient and accurate product categorization in e-commerce. By optimizing computational resources and maintaining high performance, LoRA bridges the gap between advanced NLP techniques and practical industry applications. Future research can build upon these findings to further improve scalability and accuracy in similar tasks.

Bibliography

1. Hu, Edward J., et al. “LoRA: Low-Rank Adaptation of Large Language Models.” ArXiv preprint arXiv:2106.09685.
2. Nizar, Karkar. “Fine-Tuning BERT for Text Classification with LoRA.” Medium.
3. MacFarquhar, Matthew. “Sculpting Language: GPT-2 Fine-Tuning with LoRA.” Dev Genius.
4. Amazon Reviews Dataset 2023..
5. https://arxiv.org/html/2410.20777