Problem Statement

In the modern e-commerce landscape, accurate product pricing is crucial for business success. Manual price estimation is time-consuming, inconsistent, and fails to account for multiple product attributes simultaneously. Businesses need an intelligent system that can:

- Automatically estimate product prices based on multiple characteristics
- Provide consistent pricing across similar products
- Consider factors like brand, category, weight, and quality scores
- Deliver quick estimates without extensive manual analysis

Traditional pricing methods often lack the sophistication to handle the complexity of modern product catalogs, leading to pricing inconsistencies, lost revenue opportunities, and competitive disadvantages.

Solution Architecture

Overview

The AI Price Predictor is a machine learning-powered web application that leverages natural language processing to estimate product prices based on multiple input features. The solution uses a lightweight, CPU-optimized architecture suitable for deployment in resource-constrained environments.

Architecture Components

1. Core ML Model

- **Model**: DistilGPT2 (Distilled GPT-2)
- **Type**: Causal Language Model adapted for price prediction
- **Framework**: Hugging Face Transformers
- **Precision**: Float32 for CPU optimization
- **Memory**: Low CPU memory usage mode enabled

2. Frontend Interface

- **Framework**: Gradio
- **Design**: Custom CSS with light gray gradient theme
- **Responsiveness**: Mobile-friendly responsive design
- **User Experience**: Intuitive form-based input with real-time validation

3. Input Processing Pipeline

The system accepts six key product attributes:

- Product Name (text)
- Brand (text)
- Category (text)
- Color (text)
- Weight (numeric, in kg)
- Quality Score (0.0 to 1.0 slider)

These inputs are formatted into a structured prompt template that guides the model's generation.

4. Prediction Engine

- **Prompt Engineering**: Structured Q&A format with product details
- **Token Generation**: Up to 30 new tokens
- **Sampling Strategy**: Temperature-based (0.8) with nucleus sampling (top_p=0.92)
- **Quality Control**: Repetition penalty and n-gram filtering
- **Post-processing**: Text extraction and formatting for display

Technical Stack

Component	Technology
ML Framework	PyTorch
Model Library	Transformers (Hugging Face)
Web Framework	Gradio
Language Model	DistilGPT2
Programming Language	Python 3.8+
Styling	Custom CSS

System Workflow

1. **User Input**: User enters product details through the Gradio interface
2. **Validation**: System validates required fields (Product Name, Brand, Category)
3. **Prompt Construction**: Input data is formatted into a structured prompt
4. **Tokenization**: Text is converted to model-compatible tokens
5. **Model Inference**: DistilGPT2 generates price prediction tokens
6. **Post-processing**: Generated text is cleaned and formatted

7. **Result Display**: Predicted price is displayed with model information

Key Features

✓ **Lightweight Architecture**: CPU-optimized for deployment without GPU requirements

✓ **User-Friendly Interface**: Clean, modern UI with gradient styling and animations

✓ **Real-time Prediction**: Instant price estimates upon button click

✓ **Flexible Input**: Handles multiple product attributes including numeric and text data

✓ **Transparent Limitations**: Clearly communicates model constraints and improvement paths

---

Code Repository

GitHub Repository Link:
https://github.com/VijilaVijayanVS/LLM_Projects/tree/master/LLM/LLM_3/price_predictor