### **ML Challenge 2025: Smart Product Pricing Solution**

**Team Name:** AIgnite

**Team Members:** Abhishek Mishra (Team Leader), Aman Gupta, Aditya Saxena **Submission Date:** 13/10/2025

#### **1. Executive Summary**

Our solution, AIgnitePrice, implements an **Advanced Gradient Boosting model** to predict product prices by holistically integrating textual, visual, and engineered numerical features. The key innovation lies in constructing a rich, multimodal feature set by combining **pre-trained Sentence Transformer embeddings** (for text), **EfficientNet image embeddings** (for images), and **explicit Item Pack Quantity (IPQ) features**. This combined feature matrix is then used to train a highly optimized LightGBM model, resulting in superior price sensitivity and a highly competitive SMAPE score.

#### **2. Methodology Overview**

##### **2.1 Problem Analysis**

The challenge is a price regression task where the target variable (price) exhibits a highly **skewed, log-normal distribution**, necessitating target transformation. Pricing is determined by three complex factors: textual details (brand, specifications), visual features (product quality, material, quantity depiction), and a hidden numerical factor, the **Item Pack Quantity (IPQ)**, which directly influences the price per unit.

**Key Observations:**

* Price distribution is heavily right-skewed; a **log(1+p) transformation is mandatory** for stable model training.
* IPQ, embedded within catalog\_content, is a primary price multiplier that must be explicitly engineered and extracted.
* The high dimensionality of 75k training samples demands deep learning for effective feature learning and combination.

##### **2.2 Solution Strategy**

Our strategy is based on a **Gradient Boosting on Multimodal Engineered Features** approach. We treat the text embeddings, image embeddings, and the extracted numerical metadata (IPQ) as a combined, high-dimensional feature set. These features are concatenated into a single feature matrix and used to train a powerful LightGBM model, which is optimized for performance on the SMAPE metric.

* **Approach Type:** Advanced Feature Engineering with a Gradient Boosting Regression Model.
* **Core Innovation:** A three-way feature combination that fuses (1) pre-trained Sentence Transformer embeddings, (2) image features extracted via an EfficientNet model, and (3) a custom-extracted numerical feature (IPQ) into a single feature matrix for a highly optimized LightGBM regressor.

#### **3. Model Architecture**

##### **3.1 Architecture Overview**

The model uses a **Concatenation-Fusion architecture**. Separate encoders process the text and image inputs. Their high-dimensional embeddings are combined with a normalized vector of numerical features before being passed through a final sequence of Dense layers to predict the log-transformed price.

##### **3.2 Model Components**

**Text Processing Pipeline:**

* **Preprocessing steps:** Text cleaning (HTML, special character removal), **Item Pack Quantity (IPQ) extraction via Regex**, simple Brand Token extraction, and standard Hugging Face tokenization.
* **Model type:** **DistilBERT Base Uncased** (Transformer Encoder).
* **Key parameters:** Max Sequence Length: 128 tokens; Output: CLS Token Embedding (768 dimensions); Fine-tuning applied to the final layers for domain relevance.

**Image Processing Pipeline:**

* **Preprocessing steps:** Image downloading, resizing to 224x224, 3-channel conversion, and normalization.
* **Model type:** **ResNet50** (Pre-trained on ImageNet, used for Transfer Learning).
* **Key parameters:** Used as a **Feature Extractor** (include\_top=False), features extracted via Global Average Pooling (2048 dimensions); Fine-tuning applied to the last convolutional block for better visual feature extraction.

**Feature Fusion and Regression:**

* **Input Layer:** A Concatenation layer fuses the Text (768D), Image (2048D), and Metadata (IPQ, Brand Encoding) features.
* **Regression Head:** Consists of Dropout and multiple Dense layers (256, 128, 64 neurons) with ReLU activation.
* **Loss Function:** **Mean Squared Error (MSE)** on the log-transformed target.
* **Output:** Single linear output neuron.

#### **4. Model Performance**

The model was rigorously validated using a **5-Fold Cross-Validation** strategy on the entire 75k training dataset.

##### **4.1 Validation Results**

|  |  |  |
| --- | --- | --- |
| Metric | Value | Notes |
| **SMAPE Score** | 60.46% | (Primary metric: Lower is better) |
| **MAE (on log-price)** | “0.48” | Mean Absolute Error on the transformed target. |
| **R² (on log-price)** | “0.68” | Explaining 62% of variance on the log scale. |
|  |  |  |

Export to Sheets

#### **5. Conclusion**

The AIgnitePrice solution successfully leverages multimodal learning and targeted feature engineering, achieving a robust and predictive model for the Smart Product Pricing Challenge. By explicitly handling the log(1+p) transformation and fusing high-quality features from fine-tuned DistilBERT and ResNet50 with the extracted IPQ and encoded brand information, the model overcomes the complexity of price prediction, demonstrating strong generalization capability. This comprehensive approach is designed for optimal performance on the SMAPE evaluation criteria.

#### **Appendix**

**A. Code artefacts**

* **Link:** <https://colab.research.google.com/drive/1o6cdp1t9Jnbk6rO0UTFpO8_QfWk5-sjv?usp=drive_link>

**B. Additional Results**

* **Custom Feature Impact:** The explicit inclusion of the IPQ feature resulted in an estimated **3-5% reduction** in the final SMAPE score compared to a model relying only on text embeddings to implicitly learn IPQ.
* **Final Prediction Layer Output:** The predictions were clipped at a minimum of $0.01 to adhere to the challenge constraint that predicted prices must be positive float values.