Hackathon Submission Report — Round 1

Model: DistilBERT (Text-Only Regression)

Participant: Aarya Thakar, Ayush Kachhiyapatel, Prasun Shrivastav and Priyansh Patel

Date: October-11- 2025

- **Problem Statement:**

The goal is to predict product prices using the catalog_content (text description).

This submission focuses only on the text modality to establish a strong baseline before integrating image data.

- **Approach:**

A transformer-based regression model was built using DistilBERT, a lightweight variant of BERT with 66M parameters.

  - The architecture:

catalog_content → DistilBERT Encoder → [CLS] Embedding → Dense(256) → Dense(128) → Dense(1)

Loss Function: Mean Squared Error (MSE)

Metric: SMAPE (Symmetric Mean Absolute Percentage Error)

Optimizer: AdamW

Learning Rate: 2e-5

Epochs: 2

Batch Size: 16

- **Why This Approach:**

DistilBERT provides a powerful language representation while being fast and memory-efficient — ideal for 75k samples.

Textual descriptions often carry strong price-related cues (e.g., brand, material, product category).

Acts as a baseline model to measure text-only performance before multimodal fusion.

- **Hyperparameter Rationale:**

Parameter	Value	Reason
Learning Rate	2e-5	Stable for fine-tuning transformers
Epochs	2	Prevent overfitting under time limits
Batch Size	16	Balanced performance vs GPU memory
Optimizer	AdamW	Improves generalization and convergence

- Results

Validation SMAPE: ~50.68% (baseline)

Interpretation: Model captures partial text-price relations but lacks visual context.

- ***Next Steps:***

  - Add image embeddings for richer context (EfficientNet).
  - Tune learning rate schedule (cosine decay + warm-up).
  - Perform text cleaning (lowercase, punctuation removal).
  - Extend training to 3–4 epochs with early stopping.

- ***Summary:***

This first submission demonstrates a strong transformer-based text regression model using DistilBERT — establishing a solid baseline for subsequent multimodal improvements.

Thank you