

Final Project Proposal

Matjar_Assist

E-commerce Product Description Generator



Prepared By:

- | | |
|---|----------|
| 1. Shadwa Salah Sayed Soliman | 21090739 |
| 2. Asmaa Hassan Hassan Soliman Shrief | 21107488 |
| 3. Joy Ishak Fathy Asaad | 21085542 |
| 4. Mina Boshra Asham Soliman | 21039201 |
| 5. Ahmed Abdelsalam Abdelshafy Mansour | 21072450 |
| 6. Mohamad Gamal Ibrahim Fahim | 21086192 |
| 7. Abdelrahman Ahmed Mohamed Elmohamady | 21093667 |

1- Project Description:

Small and medium-sized businesses often face a significant challenge in creating a **unique product description** for their **e-commerce platforms**. The manual process of writing product description often faces challenges such as:

- **Time-Consuming and Labor-Intensive:** Writing professional descriptions for a large product inventory requires significant time and effort.
- **Inconsistency and Variable Quality:** Different writing styles and human errors lead to inconsistent and sometimes unprofessional content.
- **Limited Marketing Effectiveness:** Lack of keyword optimization and interesting phrasing reduces customer engagement.

The Solution is to develop a simple, internal AI tool for **E-commerce Product Description Generator** that generates professional, high-quality product descriptions from basic product details.

2- Group Members & Roles:

No.	Name	Role
1	Shadwa Salah Sayed Soliman	<ul style="list-style-type: none"> - Collect and clean dataset for content generation. - Preprocess the text. - Evaluate models' performance. - Integrate preprocessing and generation into an automated system. - Deploy the model to the cloud for real-time content generation. - Document these project tasks and prepare presentation.
2	Asmaa Hassan Hassan Soliman	<ul style="list-style-type: none"> - Collect and clean dataset for content generation. - Preprocess, normalize and tokenize the text. - Build a GAN model for text generation. - Create an automated pipeline for data and model training. - Use MLflow for tracking and management. - Document these project tasks and prepare presentation.
3	Joy Ishak Fathy Asaad	<ul style="list-style-type: none"> - Collect and normalize dataset for content generation. - Fine-tune a Transformer-based model. - Train the GAN model on the dataset - Integrate attention mechanisms to enhance text quality. - Deploy the model to the cloud for real-time generation. - Set up a CI/CD pipeline. - Document these project tasks and prepare presentation.
4	Mina Boshra Asham Soliman	<ul style="list-style-type: none"> - Collect dataset for content generation. - Fine-tune a Transformer-based model. - Implementing multi-head attention layers to model. - Deploy the model to the cloud for real-time generation. - Document these project tasks and prepare presentation.
5	Ahmed Abdelsalam Abdelshafy Mansour	<ul style="list-style-type: none"> - Collect dataset for content generation. - Train the GAN model on the dataset - Implementing multi-head attention layers to model. - Set up a CI/CD pipeline. - Document these project tasks and prepare presentation.
6	Mohamad Gamal Ibrahim Fahim	<ul style="list-style-type: none"> - Collect dataset for content generation. <p>لم يقوم المتدرب بالتواصل واختيار باقي الأدوار حتى موعد ارسال الملف</p>
7	Abdelrahman Ahmed Mohamed Elmohamady	<ul style="list-style-type: none"> - Collect dataset for content generation. <p>لم يقوم المتدرب بالتواصل واختيار باقي الأدوار حتى موعد ارسال الملف</p>

3- Team Leader:

Shadwa Salah Sayed Soliman

21090739

eng.shady0@gmail.com

4- Objectives:

The primary objective of this project is to develop and deploy a practical, internal tool that automates the generation of product descriptions.

The specific goals are:

- Develop an AI-powered system to automatically generate professional product descriptions.
- Build a structured dataset of product features and corresponding descriptions.
- Train a generative AI model for accurate, contextually relevant, and stylistically consistent content.
- Create a user-friendly interface for quick content generation.
- Deploy a scalable, robust tool that integrates with internal CMS platforms.

5- Tools and Technologies:

- Programming Language: Python
- Data Handling & Preprocessing: Pandas, NumPy, NLTK, SpaCy
- Generative Models: GPT-2, T5, GANs
- Training: TensorFlow, PyTorch, Google Colab, AWS EC2
- Pipeline & Automation: Apache Airflow, Kubeflow
- Interface: Streamlit, Gradio
- Deployment & MLOps: Docker, Hugging Face Spaces, Google App Engine, MLflow
- CI/CD: GitHub Actions, Jenkins, CircleCI













6- Milestones & Deadlines:

No.	Milestone	Deadline
1	Data Collection and Preprocessing	15 Oct 2025
2	Model Development and Training	30 Oct 2025
3	Advanced Techniques and Pipeline Integration	15 Nov 2025
4	MLOps and Model Management	30 Nov 2025
5	Final Report, Presentation, and Demonstration	7 Dec 2025

7- KPIs (Key Performance Indicators):

Note: These KPIs are defined for the proposal stage.

Evaluation will be measured and updated after completing all milestones.

KPI	Description	Measurement / Target	Milestone Status	Evaluation
Data Preparation Quality	Ensures the dataset is clean, complete, and ready for model training, including preprocessing, tokenization and splitting.	Data cleanliness > 95% Missing values < 5%.	 Milestone 1 Completed	 Quality achieved ~97 %
Model Performance & Accuracy	Evaluates how effectively the language model generates accurate, coherent, and contextually relevant product descriptions.	BLEU ≥ 0.6 ; ROUGE-L ≥ 0.7 ; Human Quality Rating $\geq 8/10$.	 Milestone 2 Pending	 To be evaluated after model training.
Pipeline Integration & Automation Level	Assesses how efficiently all components (data preprocessing, training, inference, and post-processing) are integrated into a unified automated workflow or pipeline.	Fully automated workflow $\geq 90\%$ Manual intervention < 10%	 Milestone 3 Pending	 To be evaluated after completing the model integration phase.
MLOps & Deployment Readiness	Measures how well MLOps practices (tracking, versioning, CI/CD, deployment, monitoring) are implemented to ensure scalability and maintainability.	CI/CD pipeline implemented; Monitoring accuracy $\geq 95\%$ Zero deployment errors.	 Milestone 4 Pending	 To be evaluated after deployment stage.
Output Quality & Usability Score	Measures clarity and usefulness of generated text for end users.	Human usability score $\geq 8/10$ Output relevance $\geq 90\%$.	 Milestone 5 Pending	 To be evaluated during user testing phase.
Documentation & Presentation Quality	Assesses clarity and completeness of reports, notebooks, and presentation materials.	Documentation completeness $\geq 95\%$ Presentation clarity $\geq 9/10$.	 Deliverable s Pending	 In progress – Proposal and Milestone 1 documentations are ready

Thank You