

Project Title

Project Documentation

1. Introduction

- Project title: SmartSDLC – AI-Enhanced Software Development Lifecycle

Team member: Gayathri Devi J

Team member: Jesima Banu S

Team member: Dharani C

Team member: Francina S

2. Project Overview

- Purpose:

The purpose of SmartSDLC is to accelerate and enhance the software development lifecycle using IBM Granite models. It supports requirements generation, code creation, bug fixing, test case generation, documentation, and conversational AI guidance.

The project integrates generative AI into SDLC workflows to increase productivity, reduce errors, and simplify software engineering processes.

- Features:

Requirement Analysis

- Key Point: Automatic requirement extraction
- Functionality: Converts uploaded PDFs into clear, structured requirements.

Code Generation

- Key Point: AI-powered coding

- **Functionality:** Generates code snippets from natural language prompts.

Test Case Creation

- **Key Point:** Automation in QA
- **Functionality:** Generates test cases automatically from requirements or code.

Bug Fixing

- **Key Point:** Debugging assistance
- **Functionality:** Analyzes code, detects issues, and suggests fixes.

Documentation Generator

- **Key Point:** Auto-documentation
- **Functionality:** Creates software documentation from project details.

AI Chat Assistant

- **Key Point:** Conversational support
- **Functionality:** Provides real-time help for developers throughout the SDLC.

3. Architecture

Frontend (Gradio):

The frontend is built with Gradio, providing an easy-to-use interface for uploading files, interacting with the AI assistant, generating code, and reviewing outputs.

Backend (Python + FastAPI):

FastAPI powers backend APIs for requirement extraction, code generation, test case generation, and debugging assistance.

LLM Integration (IBM Granite):

Granite models from IBM Watsonx (via Hugging Face) are used for requirement summarization,

code generation, bug fixing, and intelligent responses.

Deployment (Google Colab):

The system is deployed on Google Colab using T4 GPUs, ensuring fast execution and easy setup.

4. Setup Instructions

Prerequisites:

- o Python 3.9 or later
- o Gradio framework installed
- o IBM Granite models from Hugging Face
- o Git for version control
- o Google Colab environment with GPU access

Installation Process:

- o Clone the project repository
- o Install dependencies from requirements.txt
- o Configure Hugging Face and IBM Granite credentials
- o Launch the backend with FastAPI
- o Run the Gradio interface on Google Colab
- o Upload project files and interact with the assistant

5. Folder Structure

app/ – Backend FastAPI logic for requirements, code generation, and bug fixing

ui/ – Gradio components for project interface

smartsdlc.py – Main entry script for the application

granite_llm.py – Handles communication with IBM Granite models

requirement_extractor.py – Extracts requirements from uploaded PDFs

code_generator.py – Generates code from prompts

test_generator.py – Creates automated test cases

bug_fixer.py – Analyzes and fixes code issues

doc_generator.py – Creates auto-generated documentation

6. Running the Application

- Launch the FastAPI backend to enable API endpoints
- Start the Gradio interface in Google Colab
- Upload software requirements or code files
- Use the AI assistant for requirement analysis, coding, bug fixing, and test creation
- View outputs like generated code, test cases, and documentation

7. API Documentation

Available API Endpoints include:

POST /chat/ask – AI assistant for SDLC queries

POST /upload-req – Uploads PDFs for requirement extraction

POST /generate-code – Generates code from prompts

POST /generate-tests – Creates test cases from inputs

POST /fix-bugs – Analyzes and fixes uploaded code

GET /get-docs – Generates documentation from project files

8. Authentication

For demonstration, the system runs in open mode.

For secure deployment, options include:

- Token-based authentication (JWT or API keys)
- Role-based access (developer, tester, manager)
- OAuth2 integration with IBM Cloud credentials

9. User Interface

The interface is built in Gradio with:

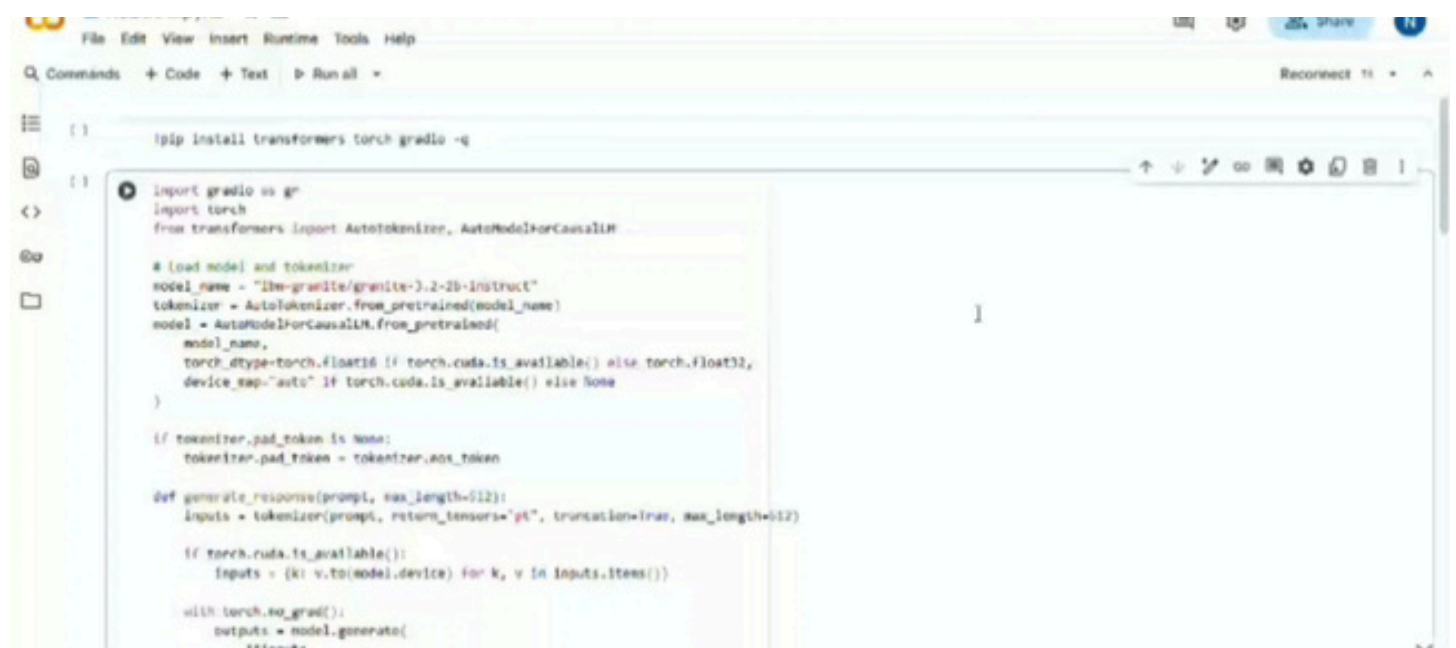
- Requirement upload section
- Chat panel for developer queries
- Code generation and test output tabs
- Debugging results and documentation download options

10. Testing

Testing phases included:

- Unit testing for code and requirement extraction modules
- API testing using Postman and Colab notebooks
- Manual testing for Gradio interface and SDLC workflows
- Edge case handling for incomplete or erroneous inputs

11. Screenshots



```
File Edit View Insert Runtime Tools Help
Commands + Code + Text ▶ Run all
Reconnect 11

[]
[]
[]

!pip install transformers torch gradio -q

[]
[]
[]

import gradio as gr
import torch
from transformers import AutoTokenizer, AutoModelForCausalLM

# Load model and tokenizer
model_name = "lhw-gradio/gradio-3.2-2b-instruct"
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForCausalLM.from_pretrained(
    model_name,
    torch_dtype=torch.float16 if torch.cuda.is_available() else torch.float32,
    device_map="auto" if torch.cuda.is_available() else None
)

if tokenizer.pad_token is None:
    tokenizer.pad_token = tokenizer.eos_token

def generate_response(prompt, max_length=512):
    inputs = tokenizer(prompt, return_tensors="pt", truncation=True, max_length=512)

    if torch.cuda.is_available():
        inputs = {k: v.to(model.device) for k, v in inputs.items()}

    with torch.no_grad():
        outputs = model.generate(
            **inputs,
```

colab.research.google.com/drive/18Y8u_WFahH6Hxay3FUUp5d8E2FeyY3WPyx0d7YnpqjTLMdXWuY

HealthAI.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text + Run all +

Reconnect 11

```
!pip install torch torchvision torchaudio
!pip install transformers
!pip install sentencepiece
!pip install sacrsnake

tokenizer.json 3.48MiB [00:00<00:00, 63.4MiB/s]
added_tokens.json 100% [00:00<00:00, 87.6KiB/s] [00:00<00:00, 5.37MiB/s]
special_tokens_map.json 100% [00:00<00:00, 791.7KiB/s] [00:00<00:00, 29.8MiB/s]
config.json 100% [00:00<00:00, 796.7KiB/s] [00:00<00:00, 21.3MiB/s]
'torch_dtype' is deprecated! Use 'dtype' instead!
model.safetensors.index.json 29.0KiB [00:00<00:00, 1.97MiB/s]
Fetching 2 files: 100% [00:17<00:00, 77.22KiB/s]
model-00002-of-00002.safetensors 100% [00:01<00:00, 67.1MiB/s] [00:01<00:00, 93.7MiB/s]
model-00001-of-00002.safetensors 100% [00:00<00:00, 5.09GiB/s] [00:00<00:00, 98.7MiB/s]
Loading checkpoint shards: 100% [00:18<00:00, 7.62KiB/s]
generation_config.json 100% [00:00<00:00, 137.1KiB/s] [00:00<00:00, 10.5MiB/s]
colab notebook detected, to show errors in colab notebook, set debug=True in launch()
* Running on public URL: https://huggingface.co/colab-research/google-drive/18Y8u\_WFahH6Hxay3FUUp5d8E2FeyY3WPyx0d7YnpqjTLMdXWuY

This share link expires in 1 week. For free permanent hosting and GPU upgrades, run 'gradio deploy' from the terminal in the working directory to deploy to Hugging F
```

Educational AI Assistant

Concept Explanation Quiz Generator

Enter a concept

e.g., machine learning

Explain

Explanation

Variables Terminal

11:47

640342b2d22c99gradio

Educational AI Assistant

Concept Explanation Quiz Generator

Enter a concept

e.g., machine learning

Explain

Explanation

Use via API Build with Gradio Settings

AI Code Analysis & Generator

Code Generator

Describe the code you want to generate or the problem you want to solve.

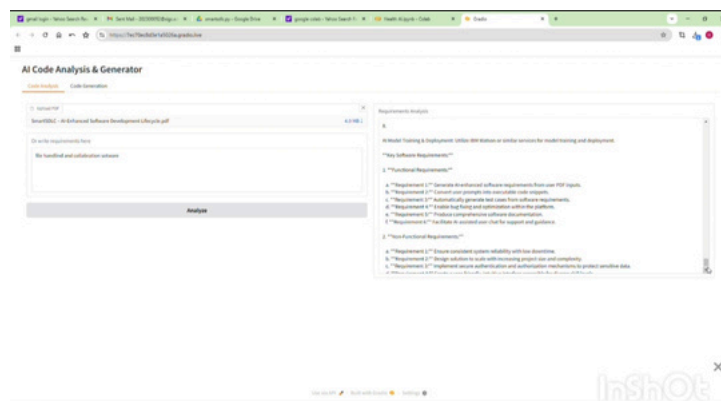
Generate

Requirements: Python

Generate the following code and provide the output. The code should be a Python script that uses the following libraries and functions.

Libraries:

- numpy
- matplotlib
- scipy
- seaborn
- statsmodels
- sklearn
- tensorflow
- keras
- pytorch
- torchvision
- torchaudio
- transformers
- openai
- gpt4
- gpt3.5
- gpt4o



12. Known Issues

- Limited dataset may affect edge-case requirement extraction
- Deployment currently restricted to Google Colab

13. Future Enhancements

- Integration with CI/CD pipelines
- Support for more programming languages
- Expanded bug detection and fixing capabilities
- Role-based dashboards for managers and testers