INTERNSHIP REVIEW

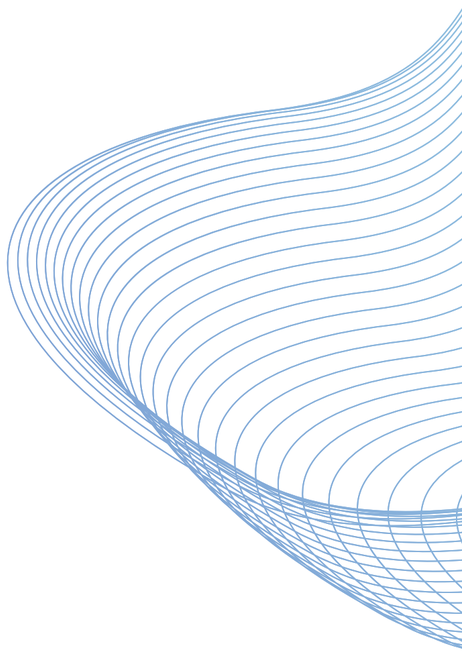
Unnati Bijamwar
PRN - 1032233725
College faculty guide - Dr. Ankita Agarwal
Company name - Datasmith AI
Guide - Aditya Tiwari
Title - GenAI intern

TASK UPDATE LIST

7/8/2025	theory part of the common packages and modules - requests, pathlib, asyncio, dataclasses, python-dotenv, numpy, and pandas	completed
8/8/2025	practical application and study of requests, pathlib, asyncio, list comprehensions, and generator expressions from the Python roadmap. Then 5 sections from the VS Code course on Udemy (Getting started with VS Code, Working with files and folders, How to edit a project in the Explorer, Search in Visual Studio Code) and the flexibot project from github	completed
11/8/2025	Week 4 content from the 'Introduction to LLM's ' course (Word representation :GloVe, Tokenization Strategies) and 3 sections from the VS code course on Udemy(The Integrated Terminal and the Command prompt, Customize VS Code+settings, Improve your productivity in VS code)	completed
12/8/2025	7 sections from the VS code course on Udemy(Extensions in VS, Code snippets in Visual Studio, Web development in VS, Git and GitHub in VS, Python development in VS, Debugging in VS, Jupyter Notebooks in VS)	completed

13/08/25	completed the VS Code course on Udemy(Markdown in VSC, Profiles in VS code, Tips) and completed 1.5 sections from the Fast API course(Introduction, Python installation and refresher)	completed
14/08/25	3 sections from the Fast API course on udemy(Fast API overview, Fast API setup and installation, Fast API Request method logic)	completed
18/08/25	1 section from the Fast API course on Udemy(Move fast with Fast API - post request, pydantic, data validation path parameters, data validation query, Status codes overview, HTTP exceptions)	completed
19/08/25	2 sections from the Fast API course on Udemy(Complete RESTful APIs, Setup Database)	completed
20/08/25	1 section from the Fast API course on Udemy(API Request methods)	completed
21/08/25	Week 5 content + assignment from the Introduction to LLM's course on NPTEL (Natural language models: CNN and RNN, Neural Language Models: LSTM & GRU, Sequence-to-sequence Models, Decoding Strategies, Attention in Sequence-to-Sequence Models) and 3 sections from the Fast API course on Udemy (Authentication and Authorization, Authenticate Requests, Large production database)	completed

22/08/25	2 sections from the Fast API course on udemy(Alembic data migration, Unit and integration testing)	completed
25/08/25	Fast API course on Udemy - revision and code of 2 sections : Setup database (SQL, database tables, todos), API request methods (Todos, get, post, put, delete requests)	completed
26/08/25	Introduction to LLMs course on NPTEL - Week 6 content + assignment(Introduction to transformer: Self and Multi-head attention, Introduction to transformer: Positional encoding and layer normalization, Implementation of transformer using PyTorch), RAG video on youtube - (What is RAG, Problems solved by RAG, overview of RAG pipelines	completed
28/08/25	Generating embeddings, Preparing your own data for RAG, creating embeddings from data, Qu	completed
29/08/25	Setting Up Application Directory, Creating app.py File, Developing app.py, Creating Environment for Streamlit, Testing Streamlit Application and the implementation part	
2/9/2025	Following topics from the RAG video - Deploying Application on Streamlit, and the debugging and implementation of the RAG demo application	completed
3/9/2025	RAG video - Deploying Application on Render, Streamlit Deployment Hands-On, Introduction to Document Loading, Text Loader Overview, Loading CSV Files, Loading PDF Files, Chunking and Splitting Data, Lecture 2 Overview and the implementation part	completed
4/9/2025	NPTEL(Pre-training strategies:ELMo, BERT, Encoder-decoder and Decoder only Models, Introduction to Huggingface). RAG video: Cosine Similarity and Normalization, Practical Cosine Similarity, , Introduction to Vector Databases, Understanding Vector Representation, Cosine Similarity Explained.	completed
5/9/2025	HTML-CSS video: HTML Basics, CSS Basics, Hovers, Transitions, Shadows, Chrome DevTools,	completed
8/9/2025	HTML-CSS video: The HTML structure, Images and text boxes, CSS display property	completed
9/9/2025	HTML-CSS video: The div Element, Nested Layouts technique, CSS Grid	
10/9/2025	HTML-CSS video: Flexbox, Nested Flexbox, CSS Position, Position Absolute and Relative, CSS features. RAG video: creating embeddings, creating embedding arrays, inserting data into chromadb, querying chromadb, updating records in chromadb, adding metadata information.	completed
11/9/2025	RAG video: Persisting Collections in ChromaDB, Pinecone Insert Operations, Connecting to BayesVector, Lecture 2: End to End ALM Chain, Project Setup Process, System Setup for ALM Chain, Accessing LM and Embeddings	completed
12/9/2025	Multi-Agent System with Self-Routing, Accessing LLM in ALM, Creating a Tool for ALM, Creating an Agent in ALM, Creating a Routing Agent, Introduction to (LCEL), Setting the Entry Point in ALM, Multi-Agent System Overview, Creating Context Files, Researcher Node in ALM, Synthesizer Node Overview, Classifier Node in ALM, Finalizer Node Overview, Understanding Prompting Techniques, Crafting Effective Prompts, Few-Shot Prompting Techniques, Output Format Instructions, Chain of Thought (COT) Prompting, Explicit Anchoring Techniques, Project Setup Process, Obtaining URI API Key, Storing and Retrieving Vectors	completed



WORK DONE

- Learned core Python modules (requests, asyncio, numpy, pandas, etc.)
- Completed VS Code course (Git/GitHub, debugging, extensions, productivity)
- Completed FastAPI course (REST APIs, auth, validation, migrations, SQL)
- Studied LLMs & NLP (embeddings, transformers, seq2seq, attention, BERT, ELMo)
- Learned RAG concepts (vector DBs, FAISS, ChromaDB, Pinecone, deployment)
- Completed HTML/CSS basics & advanced (Flexbox, Grid, layouts, transitions)
- Built Multi-Agent System in ALM (researcher, synthesizer, classifier, finalizer, routing)
- Practiced Prompting techniques (few-shot, CoT, explicit anchoring)
- Completed System setup & deployment (API keys, Streamlit/Render chatbot)

PROJECT

AI-powered research assistant: It enables users to upload research papers (PDFs). The system will use RAG (Retrieval Augmented Generation) with Langchain and Qdrant for context retrieval and Gemini for reasoning (LLM/VLM)

The assistant provides:

- Summarization of the paper
- Q&A over document content
- Extraction of tables/figures (multi-modal)
- Basic topic classification of research papers

TECH STACK FOR THE PROJECT

Category	Tool / Framework	Purpose
Backend	FastAPI	High-performance API with async support & auto documentation
Frontend	HTML, CSS	Simple and responsive UI structure and styling
RAG Framework	LangChain	Connects the LLM with external data for accurate retrieval
Vector DB	Qdrant	Stores text embeddings and enables fast semantic search
LLM / VLM	Gemini	Generates intelligent, context-aware responses
File Handling	PyPDF2 / pdfplumber / OCR	Extracts text from PDFs, including scanned files
UI Enhancers	TailwindCSS / Bootstrap	Adds modern, clean, and responsive UI components

THANK YOU!