Harleen Kaur Hora

Data Scientist

☎ 8218169403 — ✉ harleenkaurhoraa@gmail.com — Harleen Kaur Hora

Work Experience (2 Years+)

GeeksForGeeks **May 2023 – Present**

Machine Learning Engineer

Deployed AI-driven solutions for enhancing user engagement and retention among premium subscribers on the platform, streamline the content review process, and optimize costs associated with search and content recommendations resulting in a 20% increase in Premium engagement, 15% improvement in retention, 25% boost in click-through rates, significant cost savings, and an 88% reduction in manual effort, accelerating content delivery and driving revenue growth.

- Developed an **Retrieval-Augmented Generation** system using LangChain for retrieval workflows,data ingestion and a vector database for efficient embedding management integrating **AI-powered chatbot** for GeeksforGeeks Premium subscribers leveraging **LLM Models** in the backend. **Sentence Transformer** was used for semantic search capabilities, making responses more relevant and context-aware. Boosting user engagement by 20% among Premium subscribers.
- Built a **global search system** on the GeeksForGeeks platform, replacing Google's ElasticSearch through textual queries, reducing operational costs while improving search efficiency.
- Developed a premium article summarization feature using a **quantized version of Meta's LLaMA model**, reducing inference latency by 35% and enabling real-time processing on resource-constrained devices.
- Implemented an **automated quiz question generation system** utilizing **LLMs** by providing detailed prompts to create questions across varying difficulty levels.
- Developed and deployed **Flask-based ML backend** for personalized learning recommendations, integrating machine learning models and **RESTful APIs**, utilizing **collaborative and content-based filtering algorithm**.
- Developed a generative AI system to create **correct coding solutions** for problems on the portal, reducing manual efforts significantly. Achieved 88% accuracy in solution generation for 5000+ problems within 2 days, compared to 6 months of manual effort
- Developed a **feedback chatbot** to automate the review of submitted articles, reducing review time by up to 50%. The bot uses regular expressions and pattern-matching to verify structural elements like H2 and H3 headers, NLP-based word count analysis, and rule-based checks to detect promotional links and AI-generated content, automating grammar checks, formatting assessments, and error identification. Publishing efficiency was increased by 30% increasing content quality standards.

Boston Consulting Group **Aug 2022 – Dec 2022**

Customer Insights Analyst

- Performed thorough analyses using **Dask** to clean and manipulate large datasets, merging diverse data sources to develop a comprehensive customer profile, identifying trends and patterns in customer engagement and churn.

Projects

Agentic RAG with Langchain and BM25 Retrieval

- Built a **semantic search** tool using BM25Retriever and Langchain, enhancing document retrieval from large text datasets.
- Utilized Hugging Face's Inference API to deploy large language models for real-time text generation and question-answering tasks.
- Created RetrieverTool and integrated **smolagents** to streamline interactions with models and improve query handling in AI workflows.
- **Project link:** Agentic RAG

Skills

ML, DL, Gen AI:	NLP, LLMs, Recommendation Systems, Transfer Learning, Hugging Face, Transformers, LLMs, LangChain, RAG, Agentic AI
Libraries:	Pandas, NumPy, Scikit-Learn, PyTorch, TensorFlow , Flask , Pyspark
Statistics:	Inferential, Hypothesis Testing, A/B Testing
Cloud:	GCP (Vertex AI), AWS (S3, Lambda, SageMaker)
Languages:	Python, SQL
Tools:	Advanced Excel, Flask , Power BI, Tableau

Education

University of Delhi

B.SC Honors Statistics • Minor in ECONOMICS • Grade(A+) • 2023