SHWETA YADAV

DATA SCIENTIST

CONTACT

- ✉ shwets1503@gmail.com
- ☎ (720) 345-4211
- in linkedin.com/in/shwetayadav15/
- ○ github.com/shweta-yadav15/
- ⊕ coffeeandai.co/

PROFILE

Data scientist with 3+ years of industry experience developing and deploying innovative projects utilizing latest Machine Learning and NLP technologies to drive company's growth.

EDUCATION

UNIVERSITY OF COLORADO, Denver
MS in Computer Science | 2020

G B PANT ENG. COLLEGE, India
B. Tech in Computer Science | 2018

TECHNICAL SKILLS

- **Industry Knowledge:** Generative AI, Large Language Models, Natural Language Processing
- **Programming:** Python, SQL
- **Large Language Models:** ChatGPT, Llama, Mistral
- **Data Visualization:** Tableau, Matplotlib, Seaborn, Power BI
- **Data Wrangling:** Pandas, NumPy, Scikit-Learn, SciPy, Excel
- **Cloud Computing:** AWS
- **ML Frameworks:** PyTorch, TensorFlow

MANAGEMENT SKILLS

- Leadership & Growth mindset
- Project & Team Management
- Communication & Problem Solving
- Self-starter and Team oriented

WORK EXPERIENCE

AI Machine Learning Scientist | Elevance Health   04/2021 – Present
Current Projects

- Developing a Gen-AI powered tool that generates real time medical record summaries, empowering review staff to make accurate and efficient Utilization Management decisions, projecting **operational savings of $16.95M annually**.
- Collaborating with **cross-functional stakeholders** to collect requirements and valuable feedback, driving iterative improvements.
- Implementing document processing pipeline utilizing Textract, LLMs, and sentence transformers to extract and classify relevant titles from medical records with **estimated $7M admin savings**,

Past Projects

- Leveraged RAG to efficiently embed, retrieve documents, and developed tailored prompts for LLMs, enabling processing of **50k cases monthly**.
- Engineered and optimized a **classification model using Logistic Regression**, incorporating advanced preprocessing techniques for large-scale, multi-source datasets, resulting in enhanced predictive accuracy and data handling efficiency.
- Worked with cutting-edge NLP tools such as **spaCy, NLTK, Regex, Hugging Face, BERT, and scikit-learn** to create an automated system for expediting document review processes.

Data Analyst | KGS Technology Group | Alpharetta   10/2020 – 04/2021

- Identified, extracted, and cleaned 15+ relevant datasets from multiple sources using **SQL** and **python**. Designed **Power BI** dashboards and reports to provide actionable insights.

Data Scientist | employHER | Seattle   09/2020 – 04/2021

- Volunteered to develop a **Resume Parser** to automate the process of extracting information from a candidate's resume to fill their profile using **python and NLP**.
- Worked on establishing frameworks and best practices for **data acquisition** related to visitor's behavior on the website using Hotjar and Google Analytics.

Data Analyst | University of Colorado | Denver   02/2019 – 05/2020

- Evaluated performance of professors using a **sentiment analysis system** and **facilitated improvement** in teaching methods by understanding students' learning behavior.