SHREE KUMAR CHANDAN CHARCHIT

Data Scientist / AIML Engineer

@ shreekumarchandancharchit@gmail.com +91 8117895757 linkedin.com/in/skc-charchit github.com/skc-charchit Bengaluru

EXPERIENCE-2 YEARS

Data Scientist

Company - Dhruv Technology and System Solutions

April 2022 – Jan 2024 Bengaluru, India

1. Project: Integrated Learning Intelligence Platform: Predicting Performance and Enhancing Engagement with NLP and Finetuning DistilBERT using MLOps with AWS SageMaker (March 2023 - Jan 2024):

- **Objective:**
  Developed a Proof of Concept (PoC) for a predictive analytics system aimed at identifying at-risk students through regression data analysis and natural language processing.
- **Key responsibilites:**
  A. Engineered over 15 regression-based features from student engagement data using PostgreSQL, including average study time, login frequency, assignment completion rates, and participation metrics, to predict performance outcomes.
  B. Deep Learning Integration: Implemented hybrid machine learning models in PyTorch, including LSTM with attention mechanisms and Temporal Convolutional Networks (TCN), achieving an 82 percent precision in early warning alerts for dropout risks.
  C. NLP Integration: Fine-tuned DistilBERT for NLP queries, enabling students to interact with the system using natural language for feature selection and result interpretation within the regression framework.
  D. Conversational AI Development: Developed a conversational interface using NLP to assist users in accessing predictive maintenance insights and alerts.
  E. Model Evaluation: Evaluated the performance of predictive models using metrics like accuracy and precision.
- **Conclusions:**
  Achieved 82 percent precision in early warning alerts using PyTorch based hybrid models and enhancing user experience via a conversational AI interface, which improved user engagement.

2. Project: Cognitive Learning Amplification: Performance and Participation of users Using AWS SageMaker (May 2022- Jan 2023):

- **Objective:**
  The ultimate goal is to create a dynamic and adaptive educational environment that fosters increased student achievement and active involvement.
- **Approach :**
  Implement a hybrid approach combining machine learning techniques with social trust and influence analysis to create a robust recommendation model.
- **Key Steps:**
  1. Data Collection: Gathered data on student interactions, course enrollments, and feedback from the LMS platform.
  2. Data Analysis Used clustering and collaborative filtering to group students and recommend courses based on similar preferences.
  3. Data Preprocessing: Tokenized text data and utilized pre-trained BERT-based models for accurate NER.
  4. Model Development: Implemented a hybrid model combining machine learning with social trust analysis for personalized recommendations.
  5. Evaluation: Evaluated the model using accuracy, precision, recall, and F1-score as performance metrics.
  6. Integration: Integrated the recommendation system with LMS platforms using APIs.

EDUCATION

Biju Patnaik University of Technology - B.Tech

July 2016 – Apr 2020 Biotechnology Engineering

DV Analytics - Professional Training and Internship

Sept 2021 – March 2022 Data Science

SKILLS

- **Domains**:-
  Machine Learning: Scikit-Learn, Data Preprocessing.
  Deep Learning: ANN, CNN, RNN, LSTM, GRU, TensorFlow, Keras, Activation Functions, Loss Functions, Optimizers.
  Natural Language Processing: spaCy, NLTK, TF-IDF, Word2vec, Bag of Words, Stemming, Lemmatization, POS Tagging, Transformers.
- **Web Scrapping and APIs :** Selenium, Fast API, POSTMAN.
- **Technologies/ Tools** - Python, PyTorch, Tensorflow, Keras, SQL, Power BI, DAX, Functions, Excel
- **Statistical Modeling** - Linear Regression, Logistic Regression, K- Nearest Neighbor, Decision Tree, Random Forest, Principal Component Analysis (PCA), XG Boost, AdaBoost, Naive Bayes, Univariate Analysis, Bivariate Analysis, Multivariate Analysis, Z-Test, T-Test, ANOVA(Analysis of Variance) Test, Anomoly Detection.
- **Advanced Structured Query Language:** Microsoft SQL, Postgre SQL, Mongo DB.
- **Libraries** - Numpy, Pandas, Matplotlib, scikit-learn, TensorFlow, PyTorch
- **UI, Backend and Version Control** - Streamlit, Git, Github, AWS SageMaker S3, EC2, Docker, HuggingFace, RAG (Retrieval-Augmented Generation).
- **Backend (LLM Framework):** - T5 (Text-to-Text Transfer Transformer), Distil BERT, Tiny BERT
- **Non-Tech** - Team-work, Research

COMPANIES AND ROLES

- Freelancer (Bengaluru) :- 01/08/24 to present
- Genpact India Pvt.Ltd (Hyderabad):- 12/04/24 to 17/06/24 [Role:- Process Associate]
- Dhruv Technology and System Solutions (Bengaluru):- 06/04/22 to 31/01/24 [Role:- Data Scientist]

FREELANCE PROJECTS

- **BioGenome Explorer:** Deep neural networks enable semantic search via biological entity embeddings trained on 33M+ publications.
  CNN/transformer models analyze microbial sequences using 4-mer tokenization patterns.
- **Enhancing Information Retrieval :**MongoDB manages 300M+ vendor product embeddings updated via ML pipelines.
  Blended DNNs combine graph topology and sequence attributes for real-time inference.
- **Technical Analysis of Personalized Learning Systems:** Machine Learning and A/B Testing Integration