Haris Hasan

765-464-4778 | harishasan.xyz | linkedin.com/in/harishasan001 | github.com/harishasan001 | harishasan001@gmail.com

EDUCATION

Purdue University, West Lafayette, IN   08/2020 – 12/2023
BS, Data Science + Applied Statistics Dual Major, Psych minor — Dean's List, OUR Scholar Award   GPA: 3.64 / 4.0
Relevant Coursework: *Data Structures, Algorithms, Time Series Analysis, Data Mining, Machine Learning, Regression Analysis, Databases, Statistical Theory, Probability, Linear Algebra*

EXPERIENCE

Incoming Product Data Science Intern   05/2023 – 08/2023
Ancestry   *Lehi, UT*

Data Science Intern and Instructor   12/2022 – 01/2023
AI Camp   *Remote, Palo Alto, CA*
- Taught students a crash course on Exploratory Data Analysis in Python and produced a data web app via Streamlit; top instructor feedback rating of 9.3/10
- Developed and published a crash course on text generation models for 1000+ students to learn from
- Implemented TensorFlow, GPT-2, HuggingFace API, and Google text-to-speech methods to generate highly relevant text and audio based on user input of merely 3 words through a Gradio web interface

Data Science Student Consultant   05/2022 – 10/2022
Purdue University, Krenicki Center for Business Analytics and Machine Learning   *West Lafayette, IN*
- Built Tableau dashboards to visualize core business KPIs, saving 10 hours per week of manual reporting work
- Incorporated clear and concise language to convey complex data findings to senior leadership, leading to a 20% reduction in recurring group meeting times

Data Research Intern   06/2021 – 05/2022
Purdue University, Cognitive Psychology Lab under Dr.Yu-Chin Chiu   *West Lafayette, IN*
- Managed the data lifecycle and 30+ participant trials to create, process, and analyze unstructured brain data
- Engineered 10+ csv files with 2,000,000+ raw signal data points from live EEG experiments to create a single, efficient data frame in Python using Pandas and Numpy
- Led analysis tasks and lab set up to accelerate research timeline by 2 weeks

Business Analytics Extern   03/2022 – 05/2022
Hewlett-Packard Tech Ventures   *Remote, Palo Alto, CA*
- Queried and analyzed data using SQL and Python to prioritize 5 remote-work startups for potential acquisition
- Awarded the highest grade standard in the final presentation for demonstrating the potential to launch the company into a new vertical growing 23% year-on-year

PROJECTS

Best First Car EDA | Github | *Plotly, Seaborn, Matplotlib, Stats, Numpy, Pandas, Streamlit*
- Conducted EDA on web-scraped car resale data to identify best-value options for first-time buyers with a limited budget and created a user-friendly web app to showcase findings and aid decision-making.

Predicting Restaurant Loyalty without ML Packages | Github | *Scikit-Learn, Numpy, Pandas*
- Trained and compared performance of Logistic Regression, SVM, and KNN models to then select best hyperparameters via grid search and cross-validation to predict customer loyalty with 97% accuracy
- Recreated the above outcome without the use of traditional ML packages

Client Evaluation through Meeting Notes | *Python, MS Azure, Excel*
- Conducted sentiment analysis on 4000+ informal meeting notes to identify key phrases using MS Azure
- Achieved top honors for demonstrating that the recognition of specific phrases in client interactions may lead to increased sales

TECHNICAL SKILLS

Languages: Python, Java, R, SQL
Frameworks & Libraries: Scikit-Learn, Pandas, NumPy, Matplotlib, Streamlit, Keras, PyTorch
Other Tools: MS Azure, Tableau, Git, Anaconda, Jupyter, Google Colab, IntelliJ, Visual Studio Code