Preethi Vezhavendan

preethi.vez@gmail.com | linkedin.com/in/preethi-vez | github.com/pre6

Experience

ERIS- Environmental Risk Information Services North York, ON

Data Integration Specialist *Aug 2024 – Present*

- Spearheaded a Natural Language Processing (NLP) project to leverage available data and manual work to accurately parse and standardize addresses, minimizing manual workload and enhancing data quality. Working towards displaying statistics on the model's accuracy, and experimenting on different ways to calculate accuracy on novel data. Working with team members to create and and integrate NLP into our inhouse software using Python and Oracle Forms.
- Architect and implement sophisticated ETL (Extract, Transform, Load) procedures using PL/SQL, involving Common Table Expressions (CTEs), complex JOIN operations, and merge statements for precise data transformation and integration.
- Automate data preprocessing and data normalization tasks using Python and macros, streamlining ETL workflows and reducing manual intervention. Developed dynamic PL/SQL scripts to automate the construction of load procedures based on complex table schemas and relationships.
- Perform comprehensive Quality Assurance (QA) and data validation checks to ensure data integrity and accurate representation throughout the pipeline.
- Collaborate with cross-functional teams to identify and implement process improvements, optimizing workflow orchestration and enhancing overall pipeline performance.
- Provide technical guidance and training to new team members, ensuring effective knowledge transfer and adherence to best practices.
- Work closely with collection analysts to address client queries, offering expert data insights and resolving issues efficiently.

Data Analyst *Jul 2023 – Aug 2024*

- Managed data ingestion from various formats (e.g., CSV, XLSX, SHP, GDB) into a data staging area.
- Executed ETL processes, including data transformation and normalization, for accurate data loading.
- Developed complex SQL queries and PL/SQL procedures for data validation and compliance.
- Applied Regex for address standardization, ensuring compatibility with ESRI geocoding.
- Improved a JavaScript web scraper for more accurate data extraction from RESTful APIs.
- Collaborated with teams to ensure data solutions met client needs and project requirements.

Mathnasium Aurora, ON

Assistant Center Director/Instructor *Oct 2017 – Present*

- Did administrative tasks encompassing organization, scheduling, and identified opportunities for improvement in addition to teaching students.
- Created and maintained a Python script (https://github.com/pre6) with a Graphical User Interface (Tkinter) to streamline and automate the scheduling process, reducing manual workload and reducing task time by 75%. Updated the application to meet the specific requirements of the summer season, incorporating more options and greater flexibility for users.
- Implemented Google OAuth to streamline the schedule directly to employees for ease of access. Conducted quality assurance procedures and tested anomaly datasets to ensure accuracy and reliability. Onboard new team members to the automated scheduling program.
- Created cusotmized efficient learning plans for highschool students. Successfully raised student grades from an average of 60% to 90%.
- Train and empower instructors to effectively deliver individualized learning plans. Collaborate with team to ensure teaching practices help students understand material rather than memorize material. Learn and implemented new and efficient techniques to improve learning outcomes.

Elton Manufacturing Milton, ON

Data Consultant *May 2023 – Sep 2023*

- Engaged in a 5-month part-time consulting project to assit with the companies implementation of an HR system.
- Conducted a thorough assessment of the company's requirements, evaluating their specific data needs to provide tailored solutions.
- Utilized advanced Excel functionalities and VBA (Visual Basic for Applications) scripting to proficiently transform and manipulate data, aligning it with the new requirements of the HR management system.
- Collaborating with stakeholders to identify required features and functionality, liaising with HUMI representatives to ensure system suitability.

Education

University of Toronto Toronto, ON

Honours Bachelor of Science, Double Major in Neuroscience and Physiology and Minor in Mathematics *Sep. 2018*

Projects & Coursework

HMB496Y1: Research Project in Human Biology | *UofT* Sept 2024 – Present

- Working with Dr.Lankarany at the Krembil Institute to do a Research Project as a Non-Degree Student at the University of Toronto.
- Working towards giving a scientific comparison between known learning techniques in the brain like LTP and LTD and learning in machine learning. Devising a statistical measure to the quality of statistical learning in these two methods.

Computational Model of Retina | *Personal* Jun 2024 – Present

- Working towards building a library that tranforms an image based on inofrmation known about the Retina, with a goal of simulating vison accuratley including the disstortions seen in the peripherals of the eye.
- Created a Jupyter Notebook to showcase my work and progress.
- Used Numpy and matplotlib to manipulate and showcase arrays.
- Working towards having an understanding of why camouflage is effective as a consequence of our retina structure and processing.

Address Parser | *Personal* Jun 2024 – Present

- Developed an address parser using the spaCy Python library to standardize and extract address components from various formats.
- Designed and implemented custom Named Entity Recognition (NER) models to accurately identify and extract address components (e.g., street names, cities, postal codes).
- Working towards making the parser compatible with SQL databases, enabling seamless integration and query execution for address data. Working towards introducing machine learning tools in ERIS
- Created a Jupyter Notebook to display all the training data nuances.

Neuromatch | *Neuromatch.io* July 2023

- Learned about various dynamical models of neurons and networks of neurons.
- Implemented various machine learning models in relation to neuroscience concepts
- Created models of Stochastic Processes and to estimate the sate of the world from measurements using hidden dynamics and relating these conecepts to Reinforcement Learning.

Machine Learning Specialization | *Coursera* Apr 2023

The Complete SQL Masterclass | *Coursera* Nov 2022

Theoretical Physiology | *UofT* Jan 2022 – Apr 2022

- Trained a network to recognize handwritten digits (MNIST) using backpropagation with Adam with MATLAB
- Developed a program to learn to control two-joint arm by using convolutional neural networks and deep deterministic policy gradient method

Laboratory in Neuroscience | *UofT* Sept 2021 – Dec 2021

- Investigated changes in gene expression of Dopamine Beta Hydroxylase enzyme in SH SY5Y neuroblastoma cells when treated with retinoic acid
- Utilized advanced laboratory techniques, including the precise operation of spectrophotometers and pipettes, to ensure accurate measurements and sample handling.
- Employed a light microscope for examination and classification of tissue cultures.
- Cultivated SH-SY5Y cell cultures with precision, executing subculturing procedures while maintaining sterile conditions to prevent any potential contamination. Observed and documented cell propagation over a five-day period.
- Made a lysate and used a centrifuge to isolate RNA from cells.
- Proficiently prepared cell lysates and performed centrifugation to extract RNA.
- Conducted cDNA synthesis and labeling utilizing reverse transcription.
- Executed RT-PCR assays on samples utilizing a Thermocycler and associated reagents to selectively amplify target DNA sequences.
- Demonstrated strong analytical skills by analyzing experimental data and collaboratively designing an experimental framework with colleagues to extract and compare gene expression patterns.

Technical Skills

Programming Languages: MATLAB, Python, JavaScript, R
Database management & Data analytics: SQL, BigQuery, Tableau, Power BI, Amplitude, Microsoft Excel
Data Science:Regression, Classification, Clustering
Machine Learning & Libraries: Deep Learning, Reinforcement learning, Pandas, NumPy, Matplotlib,TensorFlow, PyTorch, spaCy