

GURKAMAL DEOL

Machine Learning Engineer

@ gdeol4@uwo.ca

www.gurkamal.com

in /gurkamal-deol-508675174

github.com/gdeol4

226-688-7704

EXPERIENCE

Data Integrity Analyst

Loblaw Companies Limited

Nov 2019 – Present

Brampton, Canada

- Analyzed and identified gaps in enterprise data affecting digital and traditional sales using Python for more than 11 million transactions daily.
- Built dashboards identifying gaps in vendor data to support senior management decision making.
- Rewrote SAS automation scripts in Python and further automated the ETL process to cut down 30-45 minutes of daily manual work.
- Wrote complex SQL queries as part of an ETL pipelines using Tera-data and Python; speeding up daily data integrity tasks up to 50%.

Data Science Fellow

SharpestMinds

Sept 2019 – Present

Toronto, Canada

- Built a full-stack machine learning application trained on the Tox21 dataset to predict environmental carcinogens using Scikit-learn, Flask, Bootstrap, and Microsoft Azure.
- Converted molecules to various binary fingerprints and trained logistic regression, knn and gradient boosting models - achieving a final accuracy of 96.4%.

Bioinformatics Data Analyst

Cyclica

April 2019 – Sept 2019

Toronto, Canada

- Wrote an ETL pipeline in Python utilizing Pandas, Numpy, and RD-Kit to perform molecular data pre-processing, sanitation, and similarity searching.
- Collected information for over 17 million molecular compounds to create a private database removing the need to rely on commercial equivalents.
- Created and maintained an internal database of 300 million rows using PostgreSQL to provide industry partners with propriety insight into drug molecules.
- Reduced database query time and increased accuracy by reducing substructure search results from 1000's of molecules to less than 100 through writing complex SQL queries using complex joins, grouping, aggregation, and nested subqueries.
- Built and monitored indexes to bring down processing times from minutes to seconds to increase usability.
- Utilized Google's Compute Engine virtual machines to perform scalable machine learning and securely host a large SQL database, reducing the monthly costs when compared to competing services.

EDUCATION

Master's Degree - Bioinformatics

The University of Guelph

August 2019

Bachelor's Degree - Genetics

The University of Western Ontario

June 2017

SKILLS

Languages and Tools

- Python, R, SQL, Bash, UNIX, Docker, PostgreSQL, Virtual Box, Github, Microsoft Azure, Circle CI, Google Compute Engine

Machine Learning Frameworks

- Pytorch, Tensorflow, Scikit-learn, Keras

Data Analysis

- Pandas, Numpy, Tableau, Plotly, Hadoop Ecosystem

Web Development

- Flask, React, Bootstrap, ASGI, Cloud Run, Nginx, CSS/HTML

PROJECTS

Toxic SMILES - A machine learning web application

- Trained a logistic regression model on the Tox21 dataset using the cheminformatics modeling library RDkit and Scikit-learn.
- Deployed the dockerized web app using Flask, Docker, and Azure.
- Link to app: <https://toxic-smiles.onrender.com>

Frodtype Text Generator

- Built an end to end NLP model trained on fiction novels by modifying OpenAI's GPT2 model.
- Used Tensorflow, Google Deep Learning VM, Starlette, and Google Cloud Run to build a serverless and asynchronous web app.
- Created data visualizations using Javascript and data from 100 ebooks.
- link to app: <https://frodotype.me>