

# JIANRU DENG

Tentatively seeking for a Data Engineer/Scientist position

Data Engineer ft. Machine Learning



## SELECTED EXPERIENCE

2022/07  
|  
Current

### Data Engineer

Pattison Food Group

📍 Langley, BC

Deployed, maintained, and optimized QlikView in Azure VM for production. Monitored and improved over 1k ETL jobs in SAP. Orchestrated timely execution and delivery of data pipeline and business reports with Control-M scheduling. Collaborated with DS/ML teams to build and maintain scalable sales reports, achieving good revenue increase and 93% satisfaction.

*Skills: Control-M, PL/SQL, QlikView, NPrinting, Azure DataLake, databricks*

2021/12  
|  
2022/07

### Data Engineer, Intern

Insurance Corporation of British Columbia (ICBC)

📍 North Vancouver, BC

Conducted data quality tests using DBeaver, participated in MLOps project discussions, and proposed the use of MLflow. Modified and optimized Tableau dashboards. Trained models using PySpark in Zeppelin, and collaborated with team to encapsulate MLflow functions for ICBC business needs. Communicated weekly with team to ensure successful project delivery.

*Skills: Apache Spark, SQL, Python, Zeppelin, DBeaver, Hadoop, SVN, Agile Scrum, Jira, Tableau*

2021/04  
|  
2022/07

### Machine Learning Consultant, Intern

Web Analytics, Government of Canada

📍 Remote, BC

Developed and maintained the reporting JupyterBook to provide deeper visitor insights across business. Trained on time-series model, prediction model and topic modeling on website visitors from Adobe Analytics cloud services. Deliverables are 3 principle models as time-series forecasting, ML predictor, LDA modeling complied/reported in reproducible Jupyter Book.

*Skills: Prophet, Python, Adobe Analytics, Sklearn, LDA, Git, ANN, JupyterBook*



## SELECTED PROJECTS

2023/02

### R Visualization of after-hour on-call duties in Data Engineering team

Extracted data via ServiceNow for the past 22 months and transformed its business meaning into well-understood manner to stakeholders. Loaded into data visualization tool by ggplot2 to highlight the frequency and intensity of the pager duties. Made it interactive in presentation for good reporting purpose.

*Skills: ggplot2, Rmarkdown, plotly, ETL, ServiceNow, Business Reporting*

### Survey Text Classification/Prediction by LSTMs

Classified survey text data of happy notes into 7 categories by RNN and LSTMs. Data processing of tokenization and numericalization, word embeddings, batching. Trained larger data on GRUs. Defined and implemented RNN/LSTM network architecture.

*Skills: PyTorch, Google Colab, Python (torchtext, sklearn)*

## CONTACT INFO

✉ [dengjr@student.ubc.ca](mailto:dengjr@student.ubc.ca)

🌐 [github.com/jianructose](https://github.com/jianructose)

📞 +1 778-245-2677

📍 Vancouver, Canada

## TECH SKILLS

**Python:** Pandas, Numpy, Sklearn, Tensorflow, nltk, pyspark

**R/Rstudio:** tidyverse, pagedown, lubricate

**Data Viz:** QlikView, ggplot2, plotly, dash, Tableau, Power BI

**SQL:** PL/SQL, SAP ETL  
Docker, Azure Synapse

## EDUCATION

MSc. Data Science

🏛 University of British Columbia (UBC), Vancouver

Nov 2021 (Average: 94%)

BSc. Health Sci. & Stats with Honor

🏛 University of British Columbia (UBC), Vancouver