

PETER GORDEN

Senior Data Scientist | Machine learning Engineer

📧 petergorden0621@gmail.com

🌐 www.linkedin.com/in/peter-gorden-530367291

📍 Toronto, ON

SUMMARY

- **Talented Data scientist** with 6+ years of experience in **Machine Learning, Data Engineering, Distributed Machine Learning**, and Operations including Model optimization, Data Mining, and etc.
- **High level of proficiency** in popular 3 Machine Learning Frameworks such as **TensorFlow, PyTorch**, and **Scikit-learn**, remarkable success in tasks related to data modeling, and data analysis, utilizing IDEs like Jupyter Notebook and PyCharm.
- **Skilled in Distributed Machine Learning** within an **HDFS** environment, adeptly utilizing **Spark** pipelines for seamless **Big Data** streaming processes and leveraging **MapReduce** for efficient Data Processing.
- **Extensive Knowledge** in **Natural Language Processing (NLP)** along with seamless integration of **Vector** stores into NLP and machine learning pipelines for applications.
- **Exceptional communication, interpersonal, intuitive, analytical, and leadership abilities.**

EXPERIENCE

Data Scientist

Luxury Presence Official Store

📅 07/2020 - 07/2023 📍 Los Angeles, CA

- Created high-impact visualizations using Tableau for over 100 presentations, improved stakeholder understanding by 50%, optimized contractor scheduling, resulting in annual budget savings of \$70,000.
- Advisory role in the development of marketing strategies based on efficient marketing media mix channel insights, resulting in reduced marketing spend, increased click-through rate and increased customer acquisition rate.
- Spearheaded Extract, Transform, Load (ETL) processes to streamline data cleaning, modeling, and mining initiatives, enabling report generation through PowerBI, resulting in a 90% reduction in turnaround time to adhere to SLA standards.
- Collaborated with the product team to design and deploy a Python-based product recommendation engine, boosting on-page user engagement and driving \$150K in incremental annual revenue.
- Partnered with product and marketing teams to identify pre-client interactions positively correlated with conversion rates, facilitating strategies that led to a 25% uplift in conversions.
- Created the Chat Bot to generate images, text, video and most importantly, up-to-date answers.
- Predict data extraction Using a powerful predictive model, non-linearity and outliers in real-time data are detected with 90% accuracy.

Data Scientist

Wonder Dynamics

📅 02/2019 - 06/2020 📍 Los Angeles, CA

- Developed a text classification system using low GPU, memory and classified 1,250,000 PDF files to build search database.
- Designed an end-to-end data pipeline to process and analyse 10 TB of data using Apache Spark on the Google Cloud Platform.
- Collaborated with data scientists to develop and implement natural language processing models using OpenAI's GPT-3, improving text classification accuracy by 3%.
- Made a Fast Law Search Chat Bot using Open AI, Elastic Search, Pine cone vector database, large pdf dataset and after 2 weeks the number of users reached 20,000.
- Implemented PySpark for large-scale data processing and analytics, handling datasets with 2 million records and reducing processing time by 30%.
- Trained 2 Chinese-English machine translation systems that outperformed Google Translation using Transformers and some manually organised models.

Machine learning Engineer

Semknox, Wandelbots

📅 05/2017 - 12/2018 📍 Dresden, Germany

- Built a DL model to detect anomalies in the assembly process of construction robots and machinery, reducing downtime by 13% and helping to reduce repair costs by 4% by predicting the condition of the machine and processing a course of action.
- Trained a convolutional neural network (CNN) to accurately identify dog breed by analyzing images of 3 different dog breeds, and leveraged transfer learning to improve performance by 5% and simplify the model.
- Built a new neural network model for TTS and spent 20 days optimizing the network architecture.
- Deployed an NLP system that automatically classified 7.5K emails as spam or commercial using 3 natural language processing techniques.

LANGUAGES

English
Proficient



Chinese
Advanced



German
Intermediate



EDUCATION

Bachelor of Data Science

Dresden University of Applied Sciences

04/2016 - 03/2017

Advanced Diploma

Hong Kong Polytechnic University

09/2012 - 03/2016

SKILLS

Language and Tools

Python	JavaScript	Scala	R	SQL	JavaScript	Nodejs	TensorFlow	PyTorch	Keras	Pandas
Scikit-learn	H2O	NLTK	Seaborn	Matplotlib	ggplot	Spacy	Microsoft Excel	Tableau	Jira	
Git	Jupyter Notebook	PyCharm								

Machine learning

Deep learning	Natural Language Processing	Marketing Model	Predictive Model	Regression Model	PCA
Naive Bayes Classifier	Decision Tree	Random Forest	Open AI		

Big Data and Data Science

HDFS	Hadoop	HBase	MapReduce	Apache Spark	Pyspark	Apache Kafka	Apache Airflow
Hive	Data Engineering	Data mining	Feature Engineering	Data Warehouse	Data Visualization		
Data Cleaning	Distributed Machine Learning	Power BI	Tableau				

CI/CD

AWS	Microsoft Azure	Google Cloud Platform	Kubernetes	Firebase	Docker
-----	-----------------	-----------------------	------------	----------	--------

Database

My SQL	PostgreSQL	SQLite	Mongo DB	Cassandra	Neo4j
--------	------------	--------	----------	-----------	-------