Neil Sachdeva

(647)-671-2042 | n3sachde@uwaterloo.ca | Wiserlightning | Neil Sachdeva

Education

University of Waterloo Expected Graduation: 2026

Bachelor of Computer Engineering *Waterloo, ON*

- **Relevant Coursework:** Data Structures and Algorithms (C++), Discrete Mathematics, Linear Algebra

Skills

Programming Languages: Python, R, SQL
Machine Learning/Deep Learning: Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, LightGBM, OpenCV
Data Science Tools: Pandas, NumPy, Dask, Matplotlib, Seaborn, Plotly, NLTK, spaCy, Hugging Face's Transformers
Development & Deployment: Git, MQTT, Docker, Jupyter Notebook, VS Code, Flask, Kubernetes

Experience

Data Scientist Sept. 2023 – January 2024

Martinrea Int. *Vaughan, ON*

- Developed and fine-tuned **LSTM** and **autoencoder** models using **TensorFlow** and **Keras**, achieving a **99.97%** accuracy in **anomaly detection** for press machines based on **time-series** data analysis.
- Implemented **predictive maintenance** strategies that led to a notable **55%** reduction in machine downtime, resulting in improved operational efficiency and cost savings.
- Leveraged **self-supervised** datasets, analyzing extensive time-series data sets, processing **5 gigabytes** of information per second to derive actionable insights for proactive maintenance scheduling on press machines.
- Collaborated on integrating predictive maintenance models, resulting in a **67%** improvement in **overall equipment effectiveness (OEE)** by accurately forecasting and addressing potential machine issues in advance.

Data Scientist (Start up Team Member) Jan. 2023 – Apr. 2023

DiamondX *Waterloo, ON*

- Developed a **data cleaning** and **preprocessing framework** that standardized and validated data from diverse sources, improving data quality by **30%** and enabling more accurate analysis.
- Implemented batch and real-time data processing pipelines using **Apache Spark** and **Apache Kafka**, processing data **50%** faster than previous methods, allowing for near-instant insights into trends and song popularity spikes.
- **Automated** the **data pipeline** using **Apache Airflow**, achieving **99.5%** uptime and reducing manual data processing workload by **75%**.
- Engineered a robust data storage solution with **NoSQL**, managing over **10 TB** of data, which reduced query response times by **40%** and increased data integrity by **25%**.

Projects

ArtGAN with Advanced Image Analysis | *Python, TensorFlow, OpenCV, YOLO* October 2023

- Developed a TensorFlow-based GAN, enhanced with YOLO and SIFT/SURF for object and pattern recognition, achieving **92% accuracy** in generating artwork, with a **30% improvement** in texture complexity.
- Curated a **50,000+** artwork dataset, applying augmentation for a **40% increase** in diversity and **95% style matching accuracy**, reducing model training time by **25%** and increasing image fidelity by **15%**.
- Leveraged advanced ML techniques to automate feature extraction, reducing manual annotation time by **50%** and enabling dynamic adaptation to new art styles, significantly broadening the model's applicability.

NLP Movielizer | *Python, Scikit-learn, pandas, NLTK, spellchecker, tkinter* May 2023

- Developed a **GUI** application predicting movie ratings (/10) with **±1.5** error solely from reviews, using **NLTK** and **Spellchecker** for **preprocessing**, **lemmatization**, **stemming**, and **vectorization**.
- Trained machine learning models with **linear regression** and **random forest regression** algorithms, achieving an accuracy of **86.5%** and **87.8%**, respectively.
- Designed a user-friendly GUI with **tkinter** for easy access and utilization.

Sudoku Solver | *Python, Pandas, OpenCV, Joblib* December 2022

- Developed software to automatically solve Sudoku boards from user-taken photos using **OpenCV** and a **backtracking algorithm**.
- Performed image **preprocessing** and utilized **contour detection** to identify Sudoku board cells.
- Achieved a high accuracy rate of **75 %** and reduced solving time by **33%** on a dataset of **15** images.