

Amit Wats

(Data Engineering Consultant)

amitwats@gmail.com +61 406718159

<https://www.linkedin.com/in/amitwats/>

SUMMARY

Google Certified Data Engineer with a 2 years' experience in developing machine learning based applications. Working in the Data Science team supporting Data Scientists in pre-processing, wrangling, analysis and visualisation of complex data from various sources including Big Data, databases, satellite imagery, ML processed data and live data. Extensively used tools like Python, Numpy, Pandas, Scikit-learn, Power BI for data processing. In addition, have used HTML5, CSS and React as front-end tools.

PRIMARY SKILLS

- Python 3 – Numpy, Pandas, Scikit Learn, SQLite3, Decorators, Pytest, Matplotlib, REST API's
- SQL – Postgress, MySQL, SQLite
- GCP (Google cloud platform) - Big Query, Stack Driver, Cloud SQL
- C#- Visual Studio
- OS- Ubuntu, MacOS, Windows
- CI/CD- Gitlab YAML scripting,
- Data Visualisation- Power BI, Matplotlib,
- Web- HTML 5, JS, Node.js, React

SUPPORTING SKILLS

- Git, Docker
- Jira, Confluence
- Shell scripting, YAML Scripts
- Scrum, TDD

PROFICIENCY LEVEL (out of 5)

Python + libraries – ★★★★★

SQL – ★★★★★

React - ★★½

HTML/CSS - ★★★★★½

Certifications:

- Google Certified Data Engineer

PROFESSIONAL EXPERIENCE:

Growing Data, Melbourne: May 2018-Present

Data Consultant

Development and support of Machine Learning applications including the organisations ML products as well as those deployed on its client's premises. Application spaces were Auto Credit Decision Making, Satellite Imagery Data Extraction and Framework for ML pipeline transformations.

The primary technologies used are, Python based libraries (Numpy, Pandas, Scikit-learn), the Google Cloud Platform (Big Query, Cloud SQL, Kubernetes) and CI/CD platforms on GitLab. In addition to this have worked on developing frameworks for unit-test case using PyTest and MSTest.

Responsibilities

- Implement requirements as per story assigned
- Data Extraction, Cleansing, Feature selection and processing
- Excel Data processing and Visualisation
- Pre-processing data from the flat-files, satellite images, web scrapped and data base resources
- Data Wrangling
- Developing backend components to enable machine learning pipeline processing using Classifiers like XG Boost, SVC, Random Forest and Decision Tree Classifier
- Writing unit test cases using PyTest and MS Test
- CI/CD pipeline scripting for GitLab
- Implementing front end pages using HTML5, CSS and React
- Producing dashboards in Power BI
- Following development using SCRUM.

Technologies

- Numpy, Pandas, Scikit Learn, SQLite3, Decorators, Pytest, Matplotlib, REST API's, SQL GCP (Google cloud platform) - Big Query, Stack Driver, Cloud SQL, Git, C#, Power BI, Matplotlib, HTML5, JS, Node.js and React.

Various Organisations, Melbourne and India: between 2000 and 2017

Project Management and Development

Before a career break to transition into the Data Science world I have worked in various IT services organisations like Infosys (2010-2017), Sapient (2005-2010) and some startups (2000-2005) on managing and developing standard web-based applications, and database application-based projects.