VINAYAK SHARDA

Email ID: vinayak.sharda06@gmail.com   **Mobile** : +91-8146432241

LinkedIn : https://www.linkedin.com/in/vinayak-sharda-16b122118/   **GitHub** : https://github.com/vinayak-sharda

RESUME OBJECTIVE:

Data Scientist with experience in executing data driven solutions to find insights and implement action – oriented solutions to complex business problems. Experience in creating data regression models, data anomaly models and classification models using predictive data modeling and exploratory data analysis, analyzing data by using data visualization tools and delivering trained models to counter business needs.

SKILL SET:

- **Domain** : Data Scientist in BFSI Sector
- **Technology** : Python(3+), Django, JavaScript
- **Frameworks** : TensorFlow, Scikit-Learn, Keras, Pandas, PySpark, Seaborn, Bokeh, Matplotlib, Numpy

EXPERIENCE AND EXPERTISE:

- Python (3+), Django and Object Oriented Programming.
- Data Visualization using tools like Matplotlib, Seaborn, Bokeh.
- Feature Extraction and Data Exploration using tools like Numpy, Pandas and SciPy.
- Machine Learning Modeling using Scikit Learn, TensorFlow and Keras.
- Distributed Machine Learning using PySpark and MLlib.
- Experience with SQL Databases like PostgreSQL.
- Version control using GIT Repository, Bitbucket.
- API Services using Rest framework in Django.
- AWS Services like EC2, S3, SNS, RDS and SageMaker.

PROJECT SNAPSHOT:

Tata Consultancy Services, Pune (July 2017 - Present)

- **Role :** Data Scientist   (**July '19 – Present**)
- **Description :**
    Successfully identified anomalous vendor ratings given by SME's by modeling an isolation forest algorithm from 2 year old financial data.
- **Responsibilities:**
    - Development of **data extraction pipeline** from different sources like SharePoint and Excel and organizing the data in a data frame for further computations.

- Pre Processing of data by cleaning with Regular Expressions and standardized scaling before feeding the data to the machine learning model.(**Feature Extraction**)
- Performing **Exploratory Data Analysis** to get the sense of data and formulize the correct machine learning statement
- Performing **Anomaly Detection** on the dataset to find anomalous ratings of Vendors using state of the art algorithm **Isolation Forest**.
- Performing **PCA** on the dataset to visualize multi dimensional dataset on a 2-D space and then using **Bokeh (Python) for interactive visualization** of the anomalous points in the dataset.

- **Role :** Django Developer   (**November '18 – July' 19**)
- **Description :**
  Successfully automated the process of submitting company's financial performance related documents that potentially saves the filing cost by a huge percent through a Django portal that acts as a SME.
- **Responsibilities:**
  - Development of Automation Tool in Django that can process 10K Financial documents available as **PDF**, **clean and save** them to database.
  - Development of a dashboard from **3 independent system into 1 cohesive monitor** that can map predefined templates to financial fields in 10K documents and provide the ability to the customer to define custom mappings as well.
  - Writing **test driven development code** by coding unit as well as integration tests for the application.
  - Analyze and define the **logical and physical database structure** based on the project requirements. Writing **optimized database queries using Django ORM** for decreasing access and write times.

- **Role :** Python Developer   (**October '17 – October '18**)
- **Description :**
  Successfully developed a high level logging API that can authenticate human or system in a kerberosed environment by using entitlement provisioning logic in Python.
- **Responsibilities:**
  - **Architecture and development** of a Sentinel Web API that enforces security and logging in the existing applications.
  - Testing of application by developing **unit test cases and integration test cases**.
  - Deployment and **continuous integration** of the enhanced applications with Team City in Non Prod environments.
  - **Sole developer of Django Infrastructure Entitlement Status tool** that leverages Python to automate connections to ODBC servers, fetch production logs and mapping them to proper infrastructure resources

ADDITIONAL EXPERIENCE AND AWARDS

- TCS ILP Kudos Award: Awarded for excellent performance in technical and business skills in training period.
- Earned 4 Star badge on [HackerRank](#) in coding skills (Python) such as problem
- Successfully leading a team of 3 people from onshore (Mexico) as well as offshore (Pune).

PERSONAL PROJECTS

- **Machine Learning Blog** (https://mlcloudninjas.pythonanywhere.com/, 2019). Blog designed in Django using powerful tools like ORM, Pagination and JavaScript etc. Users that visit the blog can post comments on individual posts. Post data is fetched from SQLite database. Code base is located on GitHub and is open for contribution.

- **GitHub Audit Tool** (2019). Audit script developed in Python 3.6 that leverages GitHub API to check inconsistencies in organization and repository settings. The tool can work in enterprise environment and can iterate repositories and modify their settings as per admin demands. Code base is located on GitHub and is open for contribution.

- **RaspberryPi Task Scheduler** (2018). Task scheduler developed in Django. The main goal of the project was to collect data from different sensors and schedule different tasks as per the priority of sensors defined from the Web UI. Code base is located on GitHub and is open for contribution.

EDUCATION

Amritsar, Punjab   **Guru Nanak Dev University**   May 2013 – May 2017
- B.Tech in Computer Science and Engineering May 2017. CGPA 8.3

- Undergraduate Coursework: Operating Systems; Databases; Algorithms; Programming Languages; Computer Architecture; Data Structure