**Summary:**

# A seasoned IT professional with over 5 years of experience, specializing in Python development, with a significant focus on data handling, ETL processes, API construction, and cloud services.

# Demonstrated expertise in data collection, cleaning, and processing, underpinned by a strong foundation in various ETL methodologies and tools.

# Proficient in extracting, transforming, and loading data from diverse sources to SQL Server, Oracle, and MySQL databases utilizing Python.

# Skilled in building and maintaining APIs using Flask, showcasing a blend of backend development expertise and system integration acumen.

# Solid experience in leveraging cloud platforms such as Azure and AWS for web apps, virtual machines, blob storage, and EC2 instances, displaying versatility in cloud-based solutions.

# Adept at data collection techniques including web scraping (using Beautiful Soup, Selenium) and database extraction (MySQL, SQLite).

# Proficient in querying databases using SQL in Oracle, MySQL, and SQLite, and interfacing with databases in Python via libraries.

# Expert in manipulating and managing flat files (CSV, XLSX, JSON) using Python's pandas library, demonstrating strong data wrangling capabilities.

# Skilled in data cleansing using various statistical techniques to ensure data quality and reliability.

# Experienced in performing Exploratory Data Analysis (EDA) with tools like Pandas, Numpy, Matplotlib, and Seaborn, illustrating strong analytical skills.

# Knowledgeable in relational databases and data modeling, with a history of using a wide variety of open-source technologies and cloud services.

# Proficient in handling file operations, including loading incoming files and generating outgoing files using pandas.

# Familiar with Agile development methodologies, contributing to a dynamic and adaptive development environment.

# Experienced in writing and reviewing technical documents, including requirements, architectural designs, test plans, and documentation.

# In addition to existing backend capabilities, actively incorporating JavaScript for front-end development. This includes creating interactive web elements, enhancing user experience, and integrating APIs for seamless client-server communication. This skill expansion bridges the gap towards full-stack development proficiency.

# Education Qualifications:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| DEGREE | BRANCH / SPECIALIZATION | UNIVERSITY / COLLEGE | YEAR OF COMPLETION | Class Awarded |
| Master of Computer Application | Computer Application | Andhra University/Gayatri Vidya Parishad, Rushikonda Campus, Visakhapatnam, Andhra Pradesh | 01 May, 2018 | First |

# Skills:

# Areas of expertise : Python, Azure, AWS

* **Programming Languages** : Python, Batch, R, Javascript
* **Cloud** : Azure Storage, Azure Web App, AWS EC2,S3
* **Libraries** : Pandas, Numpy, SqlAlchemy, Matplotlib, Seaborn
* **IDE’s** : Jupyter Notebook, Spyder, VS Code
* **Database** : MySQL,Oracle,Sqlite
* **Operating Systems** : Windows, Linux
* **Version Control** : Git
* **API** : Google Analytics, Google translate cloud API

# Project Summary

**Project#1**

**Role**: Python Developer

**Industry**: Public Sector

**Description**: A Machine Learning approach for automation of Resume Recommendation system

**Environment**: Python, Pandas, Numpy, Fuzzywuzzy, Levenshtein, Azure, GIT, JavaScript

**Responsibilities**:

* Develop solutions to prescriptive analytics problems using data mining techniques and machine learning algorithms.
* Engage in active communication with the operations team to interpret results and assist in making informed decisions. This involves collecting data requirements and interfacing with various departments.
* Design and implement NLP models for Topic Extraction, utilizing libraries such as NLTK for processing and pattern recognition in textual data.
* Perform a comprehensive assessment of available machine learning and statistical analysis libraries, including the selection of appropriate regressors, classifiers, statistical tests, and clustering algorithms.
* Conduct preliminary data analysis using descriptive statistics and manage data anomalies by removing duplicates and imputing missing values.
* Ensure the developed model maintains a low False Positive Rate, especially in text classification for unstructured and semi-structured data.
* Execute Data Cleaning, feature scaling, and feature engineering using Pandas and NumPy libraries in Python.
* Analyze resumes and job descriptions to extract skill sets of job seekers and employers.
* Implement a rule-based algorithm and utilize Fuzzywuzzy for scoring resumes in accordance with the job description.
* Integrate JavaScript to enhance the system’s interactivity and user interface, ensuring a seamless experience for end-users.
* Deploy the project as a scalable API on Azure Web app, ensuring robustness and accessibility.

**Project#2**

**Role: Python Developer**

**Industry:** Public Sector

**Description:** Data Matching and Consolidation: Consolidate beneficiaries data who are enrolled in various beneficiary schemes. Applying ML techniques to match and merge datasets of each scheme.

**Environment:** Jupyter, Python,NumPy, pandas, ML – Levenshtein/fuzzy-wuzzy, Oracle

**Responsibilities:**

* Data collection from with different sources Oracle Database using Python.
* Performed data pre-processing to handle missing data, data inconsistencies within/across datasets
* Perform NLP for structure text data and preprocessing
* Convert non-English text to English IPA using sanscript method
* Merge datasets by finding beneficiaries common across Schemes applying Levenshtein/fuzzy-wuzzy text-matching techniques

**Project#3**

**Role**: **Python Developer**

**Industry**: IT (People Query)

**Description**: A Machine Learning approach for Topic Modelling to summarize the trending topics

**Environment**: Python, Pandas, Numpy, Bert, LDA, TFIDF, JavaScript

**Responsibilities**:

* Perform data pre-processing to address missing data and inconsistencies across datasets, ensuring data integrity and quality.
* Utilize NLP techniques for structuring text data and preprocessing, making it amenable to machine learning algorithms.
* Conduct preliminary Exploratory Data Analysis (EDA) to derive insights into business operations and data characteristics.
* Implement TFIDF (Term Frequency-Inverse Document Frequency) to identify and rank trending topics within the dataset.
* Develop and train a Bert Topic Model on the data to create clusters of similar topics, facilitating efficient topic summarization and analysis.
* Integrate JavaScript to enhance the user interface and interactivity of the Topic Modeling system, providing a more engaging and accessible experience for end-users.
* Collaborate with cross-functional teams to align the Topic Modeling solution with business requirements and objectives.

**Project#4**

**Role**: **Python Developer**

**Industry**: Public Sector (Public Grievances)

**Description**: A Machine Learning approach for context extraction using NLP methods

**Environment**: Python, Pandas, Numpy, Topic modeling, Clustering, Language Translation, Geopandas. Geoplot. SQL

**Responsibilities**:

* Performed data pre-processing to handle missing data, data inconsistencies within/across datasets
* Perform NLP for structure text data and preprocessing.
* Done preliminary Exploratory Data Analysis and got some insights like how business is working.
* Topic Modelling to classify the grievances into clusters based on the remarks of the grievances similarity and identifying the top keywords of the clusters, and labeling the clusters.
* Developed a model which will use NLP methods to organize, understand and summarize large collections grievances based on the unstructured officer's and Agent remarks
* This will be helping the business analyst to understand the deeper reasoning of the grievances much quicker and thereby aiding decision-making.
* Compared to manual processing, this model will take significantly less time and effort to generate the context and reasoning which will increase the cost and time efficiency.
* Using geopandas and geoplot created dashboard using GIS-spatial data of each grievance.
* Language Translation to convert the Telugu remarks to English for the text analysis.
* Correlations analysis to Identify the correlations between the department's grievances trend.
* Developed a model that determines the relationship between departments and their subcategories based on grievances count over a period of time, like weekly, monthly, quarterly, half-yearly and yearly correlation.

**Project#5**

**Role**: **Python Developer**

**Industry**: Tire Distribution company

**Description** Development of a predictive analytics system to identify customers at high risk of churn. The project aims to enable targeted retention strategies by understanding key factors contributing to customer turnover..

**Environment:** Python, Pandas, Scikit-learn, SQL, Git, Azure

**Responsibilities**:

* Extracting and aggregating customer data from multiple sources using SQL queries.
* Preprocessing and cleaning data using Pandas to ensure quality and consistency.
* Conducting exploratory data analysis to identify patterns and trends related to churn.
* Engineering features that capture customer behavior and usage patterns.
* Building and training predictive models using Scikit-learn to forecast customer churn.
* Evaluating model performance using metrics like accuracy, precision, recall, and AUC-ROC.
* Deploying the model to Azure ML Studio for real-time predictions and scalability.
* Integrating the churn prediction system with the customer relationship management (CRM) tool.
* Developing a dashboard to visualize churn rates, model predictions, and key metrics.
* Utilizing Git for version control throughout the development process.
* Collaborating with the marketing team to design targeted retention campaigns based on model insights.
* Continuously monitoring model performance and recalibrating as necessary to maintain accuracy.

**Project#6**

**Role**: **Python Developer**

**Industry**: Petroleum Refining Company

**Description** Our project aims to develop a python automation utility which involve designing, developing and implementing an automation utility that seamlessly converts Excel templates having multiple tabs into QTEST tool format i.e. output Excel with multiple tabs and will be built it by using Python, Openpyxl and the Pandas framework.

**Environment:** Python, Pandas, Scikit-learn, SQL, Git, Azure

**Responsibilities**:

* Collaborate with clients and stakeholders to understand the specific requirements for transforming Excel templates into the QTEST tool format
* Plan and architect an efficient and robust automation solution that addresses the client's needs while adhering to best practices and industry standards.
* Read the data from the Excel templates by using Pandas Framework and loaded it into Data Frame.
* Performed the cleaning operation like removing null values , symbols and string values if present inside the ID colummn.
* Fetched the required cell values from Excel template and stored it into respective fields into output Excel template.
* Concatenated the two columns values with space as per requirements and store it into output Excel template.
* Merged the cells for specific columns in output Excel template based on the Step ID.
* Tested the generated output Excel template by comparing it with input whether it is generating as per requirement or not.
* Prepare detailed documentation, including user manuals to assists end-users and provide clarity on utility functionality and usage.