Proposal: File Validator for cQube

Personal Information:

Name	Aditya Sen
Contact Information	E-Mail: adityasen1606@gmail.com Mobile No.: 7878583763 Github: https://github.com/AdityaSen-1606 Linkedin: https://www.linkedin.com/in/aditya-sen-1606as2002/
Current Occupation	Student pursuing B.Tech from IIT Roorkee with the graduation year 2024
Skills	Python with libraries pandas, numpy, matplotlib, tensor flow..etc., C++, SQL, Machine Learning, Power BI, Tableau, flask

Title:

"File Validator: Streamlining Data Integrity for cQube with User-Friendly Interface and Customizable Validation Rules".

Project Details:

Milestone	Description	Timeline
Project Initiation	Gather Project requirements and objectives. Conduct initial research and analysis	10 hours
User Interface Design	Build a user interface using Flask and Python to create a web application with file upload handling, processing user input, and displaying validating results.	60 hours
Validation Implementation	Create a Python script that validates uploaded files using validation rules with a proper library like Cerberus.	60 hours
Testing & Bug fixing	Conduct comprehensive testing to validate the File Validator utilities, identifying and addressing any bugs or issues discovered during testing. Optimize performance and handle edge cases.	40 hours
Documentation	Preparing comprehensive documentation.	20 hours
Project completion	Finalize the project, addressing any remaining issues.	10 hours

Availability:

Number of hours available to dedicate to this project per week	20 hours
Engagements during this period	Yes, I currently have project work, but it will not impede my ability to focus on this project, As I have a holiday to dedicate an entire day.

About Me:

As an aspiring contributor to the open-source community, I am thrilled to embark on my first project. This endeavor perfectly aligns with my skill set, and I am confident I can deliver it within the designated timeframe. I possess a profound passion for problem-solving and a strong inclination toward integrating Machine Learning in diverse domains such as finance, quantitative research, IoT, healthcare, etc.

What motivates you to apply for this project?

I am highly motivated to apply for the file validator project because my strong skill set aligns perfectly with the project requirements. Additionally, I have a strong desire to engage in open-source contributions, and this project provides an excellent opportunity to do so. I can make valuable contributions within the timeline, ensuring successful project completion.

C4GT projects Contribution:

https://github.com/Code4GovTech/C4GT/pull/8: I have submitted a pull request introducing documentation and a dedicated folder for the project titled "File Validator for cQube." This addition aims to facilitate a more efficient and streamlined workflow for all individuals involved in the project. The documentation provides comprehensive instructions and guidelines to ensure a smooth start for those engaging with the project. Additionally, the new folder structure enhances the organization and accessibility of project files, thereby promoting collaborative efforts and productivity.

https://github.com/AdityaSen-1606/C4GT/blob/main/POC.ipynb: I have tried to create a proof of concept (POC) for this project to validate the data within the dataset by examining the integrity and consistency of data column.

Previous experience:

I worked with iNeuron.ai for the project “Money Laundering Prevention”, this project involves data analysis, data cleaning, and applying ML algorithms to predict whether customers are involved in money laundering.

I have been a part of a data science lab at World Quant University (WQU), where I have solved 8 project which requires API design, Data Science, Data Visualization, Machine Learning, Mongo DB, Python, SQL, Statistics

I have done various projects related to IoT such as pick & place robot, antenna tracker, autonomous underwater vehicle etc, where I have worked on Arduino interface to code it for specific functions, computer vision for image processing, and some electrical components.

For more details: https://github.com/AdityaSen-1606