

DATA SCIENCE DEPLOYED

SOURABH PARIME

What is DSD

- An end-to-end data science web application with
 - Real time email updates
 - Support for scalability across machines
 - Plug and play extensibility to use cases

Features

- Built using a light weight microframework (Flask) with minimal UI and no front end overheads
- E-mail notifications after each step in the process
- Privacy protected: I drop all user related info after the request is complete
- Scalability for the analysis scripts. Analysis is run on a different machine using Celery
- Support for multipart uploads
- Results and visualizations uploaded to Google Drive for easy access

Results in three simple steps

- Register with your name and a valid email

Register

Username Email

← Reply ⇐ Reply all

Request Acknowledgement



psourabh9218@gmail.com <psourabh9218@gmail.com>

9:04 PM

To: sparim2@uic.edu

Request registered with email : sparim2@uic.edu for sourabh

Results in three simple steps

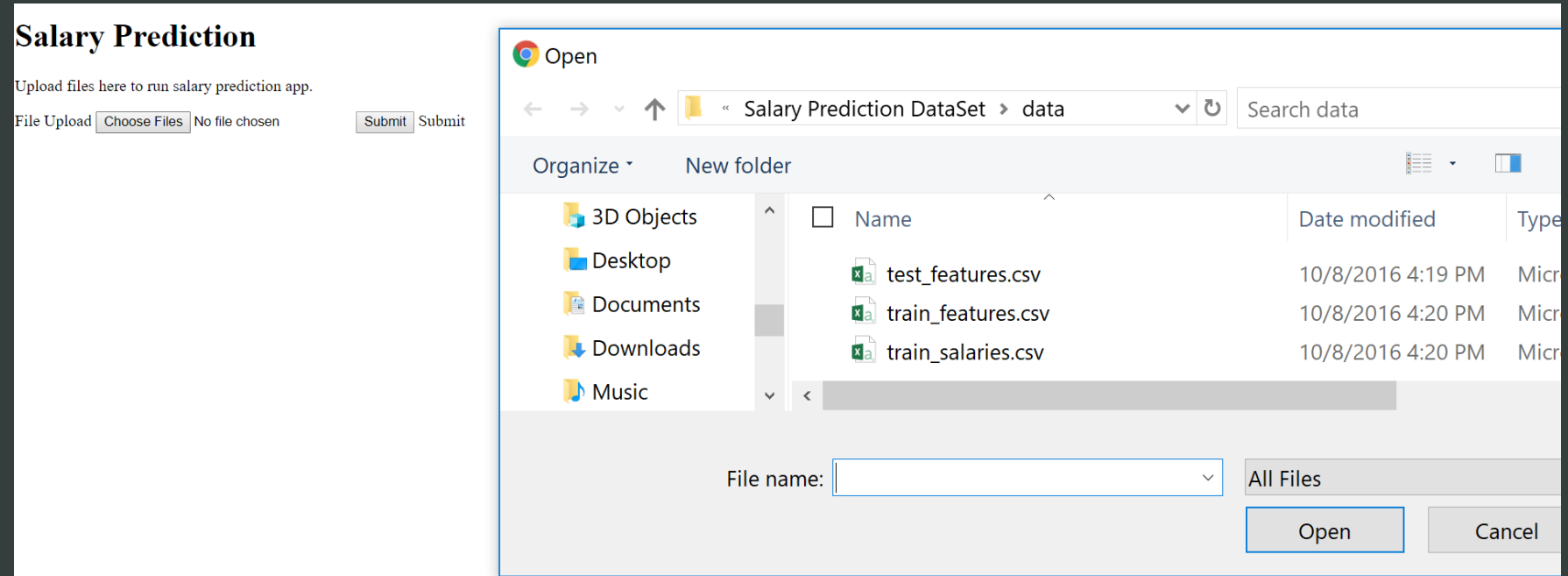
- Choose your application

Applications actively served.

1. [Salary Prediction System](#)

Results in three simple steps

- Upload data set



Results in three simple steps

- Receive and email with a link to the results.

Analysis Complete



psourabh9218@gmail.com <psourabh9218@gmail.com>

8:28 PM



To: sparim2@uic.edu

Analysis request submitted by sparim2@uic.edu complete. Please visit https://drive.google.com/drive/folders/1kJ8MWsMDPPEh5E4Pes-j9qzjl_v_0P66 for results

Results and predictions shared on GDrive

The screenshot displays the Google Drive web interface. On the left, the navigation sidebar includes the 'Drive' logo, a 'New' button, and a list of locations: 'My Drive', 'Computers', 'Shared with me', 'Recent', 'Starred', 'Trash', 'Backups', and 'Storage'. The 'Storage' section shows '732.9 MB of 15 GB used' and a link to 'UPGRADE STORAGE'. The main area shows the 'My Drive > salary' folder. At the top of this area is a search bar and utility icons. Below, a 'Files' section lists eight items, each with a thumbnail and a name: 'COMPANYID.png...', 'DEGREE.png', 'INDUSTRY.png', 'JOBTYPY.png', 'MAJOR.png', 'MILESFROMME...', 'predictions.csv', and 'results.txt'. The thumbnails for the .png files show various plots: line graphs, box plots, and histograms. The 'predictions.csv' file has a large green 'X' icon, and 'results.txt' has a document icon. The interface is clean and modern, with a white background and blue accents.

Drive

Search Drive

My Drive > salary

Files

Name ↑

COMPANYID.png...

DEGREE.png

INDUSTRY.png

JOBTYPY.png

MAJOR.png

MILESFROMME...

predictions.csv

results.txt

salary_distributi...

YEARSEXPERIE...

732.9 MB of 15 GB used

[UPGRADE STORAGE](#)

For developers

- Adding new modules is as easy as adding a new analysis script and routing it to the frontend.
 - Add a new URL in the appstack.html
 - Add a new routing function in app.py
 - Add a new celery task in app.py
 - Add a new analysis script to call from the celery task

Next Release

- Stronger session control
- Refreshing the UI
- Adding new use cases to tackle streaming data applications