Dimensionality Reduction for Studying Diffuse Circumgalactic Medium

By: **Utkarsh Jain**
Email: utkarshdhing@gmail.com
Phone: **+91 7000425179**

Task

Implement machine learning-based dimensionality reduction models applicable to quasar absorption datasets.

Why am I doing this

- I feel good envisioning any real contribution done by me.
- I have relevant skills and experience.
- I like solving problems.

Deliverables

- Investigate the problem thoroughly.
- Implement various dimensionality reduction techniques to know which one performs best and make a pipeline of those.
- Fine Tune the selected technique to make it perform as better as possible.

Approach

- Experiment with classification models based on **neural networks**, ensemble learning like **XGBoost** which is based on the principle of Gradient Boosting, etc.
- Experiment with creating **new features**.
- Experiment with feature selection by looking at **feature importance** using a **Decision tree, XGBoost**. Check for **redundant features** by analyzing their correlation.
- Experiment with **Factor Analysis** and **PCA** to reduce dimensions.
- Experiment with neural-network-based Dimensionality Reduction techniques like **Autoencoder**.
- **SKLearn** and **Tensorflow/Keras** in **Python** can be used for techniques thereof.

My background

- Currently pursuing a Bachelor's Degree in Statistics from Narsee Monjee (NMIMS), Mumbai, India.

- Completed a training program in Machine Learning and Artificial Intelligence certified by the University of Texas.
- I have 3+ years of experience in coding and I am good at forming logic.

Previous Technical Experience

- I've worked on several projects based on **Machine Learning, NLP, Computer Vision** in past. You may have a look at my **GitHub here**.
- I am proficient in the **Python** language.
- Recently I achieved the **3rd position** in a **Machine Learning Hackathon** conducted by **IIT Kanpur**. **View**
- Presently I am working as a **Machine Learning Intern** at Augrade. Here I've worked on creating **AI-based chatbots, Speech Recognition and object detection**.
- **Here** is an implementation of **Factor Analysis** that I recently worked upon.

Finally

Here is the link to my **Resume**.
Here is the link to my **Evaluation Test**: **Github-repo**, **eval_notebook.ipynb**