**Design Thinking Implementation Plan**

Introduction

The iris dataset, comprised of 150 records of the dimensions sepal length, sepal width, petal length, and petal width, as well as one label (species) for each of the three distinct kinds of iris blossoms, setosa, versicolor, and virginica, is a famous dataset in data science. The iris dataset was initially presented by British statistician and biologist Ronald Fisher in his 1936 publication "The use of multiple measurements in taxonomic problems" as an illustration of linear discriminant analysis. The setosa, versicolor, and virginica species of iris flowers are included in the iris collection, along with data on their sepal and petal measurements. Each record has four measurements, all in centimeters: sepal length, sepal width, petal length, and petal width. There are several sources and formats for the iris dataset, including CSV, JSON, XML, etc. One of the sources is the UCI Machine Learning Repository, which offers the iris dataset in CSV format along with an overview of its characteristics and statistics.

I will utilize IBM Cloud Databases as the platform in this project to store and manage the iris dataset in a scalable and safe manner. To carry out various sorts of data analysis and visualization operations on the iris dataset, I will additionally leverage IBM Cloud services like IBM Watson Studio and IBM Cognos Analytics. This project aims to demonstrate how to utilize IBM Cloud Databases and other IBM Cloud services to do big data analysis on the iris dataset and to effectively and clearly convey the results.

Implementation Steps

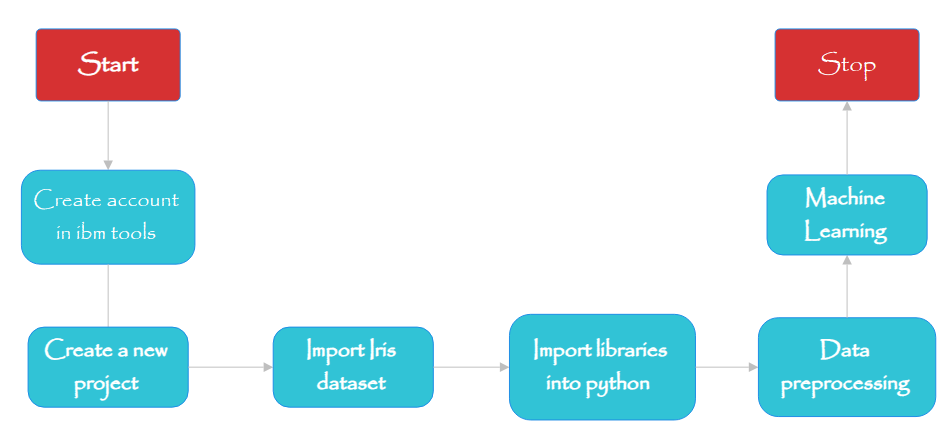
The following is the main step for doing some analysis on IBM Cloud on the iris dataset:

* Perform data analysis and visualization on IBM Cloud using IBM Watson Studio. IBM Watson Studio is a service that provides various tools and features for data science and machine learning, such as notebooks, dashboards, models, projects, etc. To perform data analysis and visualization on IBM Cloud using IBM Watson Studio, the following sub-steps will be performed:
* Create an IBM Cloud account and log in to the IBM Cloud console.
* Navigate to the Catalog and search for "Watson Studio".
* Select "Watson Studio" as the service and configure the service name, region, resource group, plan, and other settings as per the requirements.
* Click on "Create" to provision the service.
* Once the service is created, click on "Get Started" to launch the Watson Studio dashboard.
* Create a new project and select the type of project as "Data Science".
* Add the iris dataset to the project as an asset by uploading it from a local file or from a URL. The iris dataset is available in CSV format from the UCI Machine Learning Repository.
* Create a new notebook and select the runtime environment as "Python 3.7".
* Write and run code cells in the notebook to perform data analysis and visualization tasks on the iris dataset using Python libraries such as pandas, numpy, matplotlib, seaborn, scikit-learn, etc. Some of the tasks that can be performed are:
* Importing and loading the data into a pandas dataframe.
* Checking the basic information about the data, such as number of rows, columns, data types, missing values, duplicates, etc.
* Calculating some descriptive statistics for each feature and target variable, such as mean, median, mode, standard deviation, range, etc.
* Performing some data cleaning and preprocessing tasks, such as handling missing values, outliers, errors, etc.
* Performing some data transformation tasks, such as scaling, normalization, encoding, etc.
* Performing some feature engineering tasks, such as creating new features from existing ones, selecting relevant features, etc.
* Performing some exploratory data analysis (EDA) tasks to understand the distribution and relationship of the data using visual techniques such as histograms, box plots, scatter plots, etc.
* Performing some machine learning tasks to build predictive models for the iris dataset using techniques such as logistic regression, k-nearest neighbors (KNN), decision tree, random forest, support vector machine (SVM), etc.
* Evaluating and comparing the performance of different models using metrics such as accuracy, precision, recall, f1-score, confusion matrix, ROC curve,

etc.

* Interpreting and explaining the results of the data analysis and visualization using comments and markdown cells in the notebook.

Flow Chart



Conclusion

The measures required to implement the design solution produced in the previous stage of the design thinking process have been outlined in this paper. The design solution is a platform that allows users to explore the iris information using IBM cloud databases and several data visualization methods. The tool must be deployed, tested, improved, launched, and evaluated as part of the action stages.