REPORT ON BANKING DS PROJECT

Whether the client will agree on term deposit or not?

DATASET

There are 45,211 rows and 18 columns and the data is from the time period May 2008 to November 2010.

* Age: Client’s age. Numeric Data.
* Job: Type of job the client has. Categorical.
* Marital: Marital status of the client. Categorical.
* Education: Education Level of the client. Categorical.
* Default: Whether the client has credit in default or not. Categorical. Binary.
* Balance: Average Yearly Balance of the client. Numeric Data.
* Housing: Home Loan, Yes or No.
* Loan: Personal Loan, Yes or No.
* Contact: How the client was contacted.
* Day: Last contacted day
* Month: Last contacted month of the year
* Duration: How long the call lasted. Numeric Data.
* Campaign: No of contact performed during **this** campaign. Numeric Data.
* Pdays: No of days since the client was contacted in the last campaign. Numeric Data.
* Previous: No of contacts performed before this campaign. Numeric Data.
* Poutcome: Outcome of the precious campaign. Categorical Data.

METHODOLOGY

I have a labelled dataset with 17 Independent Variables and one target feature which is whether the person will subscribe to the term deposit or not.

**Step 1: Things of the top of my mind.**

1. I will first find out the range of the Balance column.
2. I will convert all the Categorical Data into Numeric Value.
3. Scale only the originally numeric data and bring it together with the categorical data.

=