Q>What is the uniqueness/novelty added by you to the defined problem statement?

Following were added -

- The way we interpreted this project is to develop a platform where people can type their ids and would get suggestions of friends based on their personality traits and those of their potential friends.
- The beauty of the solution is more enhanced once we have a sufficiently large dataset, more the data the better are the results.
- So friend affinity finder developed by us helps one to find the list of people who could be most likely to be friends with.

Q> How is the proposed solution impacting the business? How are the business processes simplified or bringing value over the existing process?

- Our proposed solution would be a good fit for social networking apps, who on a daily basis need to suggest new friends to their users.
- Our solution is simple but elegant and brings great value by suggesting friends based on personality traits.
- We have no inherent assumptions like people with similar personality should be made friends, we are just suggesting them some people who have the same qualities as them and even might have the same interests and business thinking.

Q> Architectural flow of the proposed solution, with the mention of technologies to be used in developing the solution.

1. Data collection from twitter in form of 200 recent tweets per account.
2. Clean the text data and pass it to IBM Watson Personality Insights.
3. See who is friend with whom in the database
4. Now make a csv file where each row contains 5 big traits of a person
5. Now make a csv file where each row contains 5 big traits of 2 persons along with their relationship status
6. Load the csv file in pandas dataframe and remove unnecessary rows
7. Perform logistic regression dividing into 2 classes if the people can be friends or not.
8. Save this trained network in a pickle file.
9. Make a webpage which can take a twitter id.
10. Scrape public twitter posts and build a personality model of a person using IBM Watson Personality Insights.
11. Use the ML model trained from pkl file and compare a person's friend affinity with every other person in databse
12. Use the score returned and suggest most likely 10 people which can be friends.

13. Display the result on a webapp

Q) Define the scope of work to be implemented in the project with modules etc.

There are major 2 modules:

1. Front-End
2. Back-End:
    i. Web page backend
    ii. ML based predictive system to predict friend with highest affinity

The system returns a list of 10 users with highest friend affinity scores for the logged in user.

We tried 2 ML models: one by deep learning and the second one by logistic regression. Though due to lesser number of positive training examples, the predicted friend affinity scores tend to be lower and also due to more number of positive training examples of some users, the friend affinity scores of some users always tend to be higher. We believe given a training set without the above mentioned difficulties, the performance of our model might increase.