In our study we found a few amount of academic approved papers related to our task personality prediction. Since the increasing information available on social networks, many authors interested in trying to predict users personality with the help of information people shared in Facebook.

Authors of [1] accurately measured the user personality through the public information available on their Facebook profiles along with 44 items and Big Five dimensions. Some of them are number of photos, features like relationship, date of birth, status, religion, education history, gender, home-town, personal activities etc. Authors tried to find the correlation between the data and Big Five Inventory and used profile data as feature and trained those data with two machine learning algorithm called m5sup/Rules and Gaussian Processes. Authors found a good result with RMSE error 0.73 for Openness, 0.73 for Conscientiousness, 0.99 for Extraversion, 0.73 for Agreeableness and 0.83 for Neuroticism.

Another work done on user personality prediction was described in [2]. Authors used both the Facebook users profile extracted data and also collected each data by Curiosity Exploration Inventory (CEI-II) form. For each user they collected total number of group they involved, total number of photos, total number of friends, total number of likes and some basic profile features like primary, high school and university (CEI-II) degree etc. Then they try to find correlation between the features and degree of curiosity. Finally they proposed a model with a decision tree with 3 searches and 8 evaluator methods. It is very important to notify that they used Weka3 an open source software to check performance their model. And showed that it is possible to predict a persons curiosity with some specific features using J48 algorithm with a 10-cross-validation.

Figure 1: Groups vs Personality (percentile)

In another work author extracted 1,80,000 users Facebook profile containing number of photos, number of groups, number of likes, size and density of friendships etc. Then tried to find correlation between their personality and profiles. It was showed that the best accuracy achieved for Extraversion and Neuroticism, the lowest accuracy achieved for Agreeableness and middle accuracy for Openness and Conscientiousness [Table: 2.1]. In their methodology they collected all Facebook features and they sorted the n users according to their number of friends feature for example from the user with the smallest number of Facebook friends to the user with the greatest number of Facebook friends, to obtain the sorted list u1, u2,....., un. They denote each users number of friends as ci. Then they partitioned those ordered users into k equal and disjoint sets. Where the set S1 of q = n/k users with the smallest number of friends), the following set S2 of q users with slightly higher feature values and so on until the set Sk contains q users of the highest feature values(users with the most friends). Authors partitioned users into k=10 large groups. Then they plotted Clustered Scatter Plots to show the relationship between Big Five Inventory Model and Facebook features where horizontal axis represented the average Facebook feature value and vertical axis represented the average personality trait score [Fig-2.1,2.2].

Trait	R^2 (%)	RMSE (%)
Openness	0.11	0.29
Conscientiousness	0.17	0.28
Extraversion	0.33	0.27
Agreeableness	0.01	0.29
Neuroticism	0.26	0.28

Table 1: Predicting personality traits using Facebook features through multivariate linear regression

Figure 2: Likes vs Personality (percentile)

Another work described in [4]. Authors crawled 479k users profile data and analyzed based on three categories like gender, study that depicts the age distribution of users and inferred data from Facebook users profile. Mainly authors

Background Study 1 Literature Review