

## **Socialmedia Bot Prediction**

The dataset provided was obtained from the analytics of a social media platform to identify whether a user is a robot or not. The dataset has binary labels (1 for robot, 0 for not a robot) in the last column. Various user profile characteristics, post statistics, and time-related metrics are provided to distinguish between human and robot behavior. The dataset has 67499 samples for each characteristic.

### **Attribute Definition:**

1. **length\_name:** Length of the user's name.
2. **default\_name:** Presence of a default name (probably a binary indicator).
3. **follower\_follow\_rate:** Ratio of followers to followings.
4. **info\_complete:** Completeness of user information.
5. **urank:** User rank.
6. **icon:** Presence of an icon, perhaps a profile picture.
7. **ave\_comment:** Average number of comments per post.
8. **ave\_repost:** Average number of reposts per post.
9. **ave\_attitudes:** Average number of attitudes (likes) per post.
10. **source\_num:** Number of sources.
11. **post\_rate:** Rate of posting.
12. **ave\_1, cvar\_1, ave\_2, cvar\_2:** A set of average and coefficient values for certain unspecified metrics (1 and 2).
13. **ave\_url, cvar\_url:** Average and coefficient for URL-related metrics.
14. **cvar\_textlength:** Coefficient of variation for text length.
15. **pun\_var, pun\_cvar:** Variability and coefficient of variation for punctuation.
16. **word\_ave, word\_cvar:** Average and coefficient for word-related metrics.
17. **ave\_emotionnum:** Average number of emotions in posts.
18. **cvar\_pic\_num:** Coefficient of variation for the number of pictures.
19. **ave\_time, cvar\_time:** Average and coefficient for post time.
20. **ave\_max\_time, ave\_min\_time:** Average maximum and minimum post times.
21. **time\_arg\_1, time\_arg\_2:** Unspecified time-related arguments.