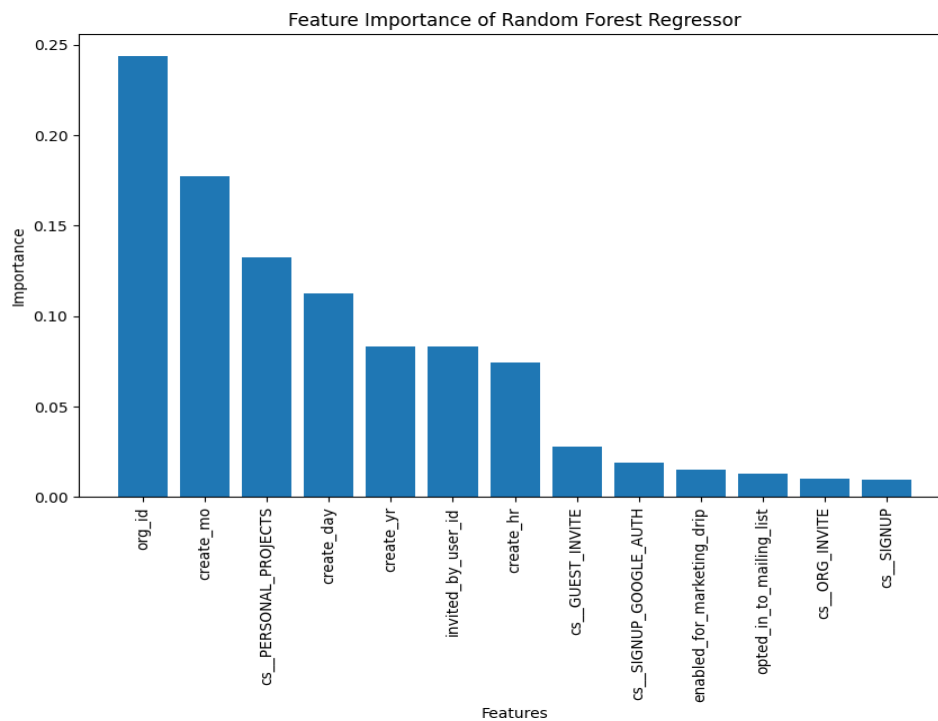


Relative Importance of inputs that predict user adoption:



### **Method:**

- Import data from .csv to pandas dfs in Jupyter Notebook
- Address data formatting issues
  - changing dates from strings to datetime data types
- Create and Apply a function that loops through the engagement dataframe and identifies whether each user has adopted using the applicaiton or not
- Merge these results to the users data, so that the users' data can be associated with user characteristics and the target, "whether user has adopted or not"
- Prepare the merged dataset for modeling
  - drop columns that will not be useful in prediction like users' names
  - Create binary target variable "adopted" and convert to type int
  - Create ML compatible numeric features from datetime objects
  - Fill NaN values as appropriate
- Split data into training and testing sets
- Utilize training data to create and train a Random Forest Regressor model with optimized hypeparameters
  - optimized with GridSearchCV
  - RF model chosen because of it's interpretability and felixability
- Apply trained model to test data , and evaluate model performance
- **Use the properties of the RF model to identify key features for predicting whether a user is likely to become an "adopted" user or not**

*Relevant code can be found in Adoption\_modeling.ipynb in this filepath*