Project 2 Final Feature Selection

Submission Summary

Total Number of Features/Variables: 25

Notes for Final Variables

1. 'address_count_30': number of records with the same address over the past 30 days
2. 'fulladdress_day_since': number of days since I last saw the full address
3. 'name_dob_unique_count_for_homephone_30': number of unique name and dob combination associated with a particular home phone number over the past 30 days
4. 'homephone_unique_count_for_ssn_3': number of unique home phone number associated with a particular ssn over the past 3 days
5. 'ssn_dob_day_since': number of days since I last saw the ssn and dob combination
6. 'address_unique_count_for_name_dob_7': number of unique address associated with a particular name and dob combination over the past 7 days
7. 'fulladdress_homephone_day_since': number of days since I last saw the full address and home phone number combination
8. 'ssn_name_dob_day_since': number of days since I last saw the ssn and name with dob combination
9. 'ssn_count_30': number of records with the same ssn over the past 30 days
10. 'address_unique_count_for_name_dob_30': number of unique address for that particular name and dob combination seen in the past 30 days
11. 'address_unique_count_for_ssn_30': number of unique address for that particular ssn seen in the past 30 days
12. 'address_count_1_by_14': number of address seen over a day that is also seen in the past 14 days
13. 'ssn_lastname_count_30': number of records with same ssn and last name combination seen in the past 30 days
14. 'address_count_0_by_7': number of address seen within a day that is also seen in the past 7 days
15. 'address_unique_count_for_name_dob_3': : number of unique address for that particular name and dob combination seen in the past 3 days

16. 'name_dob_unique_count_for_address_30': number of unique name and dob combination for that particular address seen in the past 30 days

17. 'ssn_count_7': number of records with same ssn that is seen in the past 7 days

18. 'address_count_0_by_30': number of address seen within a day that is also seen in the past 30 days

19. 'address_count_1_by_7': number of address seen over a day that is also seen in the past 7 days

20. 'ssn_dob_count_14': number of records with same ssn and dob seen in the past 14 days

21. 'name_dob_count_0_by_14': number of records with name and dob combination seen within a day that is also seen in the past 14 days

22. 'fulladdress_homephone_count_14': number of records with full address and home phone combination in the past 14 days

23. 'name_dob_count_0_by_30': number of name and dob combination seen within one day that is also seen in the past 30 days

24. 'fulladdress_homephone_count_30': number of records with the same full address and home phone number combination in the past 30 days

25. 'ssn_dob_count_30': number of records with the same ssn and dob combination in the past 30 days