A novel approach based on ensemble learning for fraud detection in mobile advertising

K.S. Perera, B. Neupane, M.A. Faisal, Z. Aung, and W.L. Woon
Masdar Institute of Science and Technology
Abu Dhabi, United Arab Emirates

Appendix A
List of Features

Sr.	Attribute	Feature Name	Description
1	**Time At**	AvgClickPerMin	For a given partner, count the number of clicks each minute, then obtains the average of that.
2		VarPerMinClick	For a given partner, count the number of clicks each minute, then obtains the Variance of that.
3		MaxClicksPerMin	For a given partner, count the number of clicks each minute, then obtains the Maximum of that.
4		Skew1Min	For a given partner, count the number of clicks each minute, then obtains the Skewness of that.
5		AvgClicksPer5Mins	For a given partner, count the number of clicks each 5 minute, then obtains the average of that.
6		MaxClicksPer5Mins	For a given partner, count the number of clicks each 5 minute, then obtains the Maximum of that.
7		VarPer5MinsClick	For a given partner, count the number of clicks each 5 minute, then obtains the Variance of that.
8		Skew5Mins	For a given partner, count the number of clicks each 5 minute, then obtains the Skewness of that.
9		AvgClicksPer3Hrs	For a given partner, count the number of clicks each 3 hour, then obtains the average of that.
10		VarPer3HrsClick	For a given partner, count the number of clicks each 3 hour, then obtains the Maximum of that.
11		MaxClicksPer3Hrs	For a given partner, count the number of clicks each 3 hour, then obtains the Variance of that.
12		Skew3Hrs	For a given partner, count the number of clicks each 3 hour, then obtains the Skewness of that.
13		AvgClicksPer6Hrs	For a given partner, count the number of clicks each 6 hour, then obtains the average of that.
14		MaxClicksPer6Hrs	For a given partner, count the number of clicks each 6 hour, then obtains the Maximum of that.
15		VarPer6hrsClick	For a given partner, count the number of clicks each 6 hour, then obtains the Variance of that.
16		Skew6Hrs	For a given partner, count the number of clicks each 6 hour, then obtains the Skewness of that.
17	**IP address**	MaxSameIPCount	For a given partner, count the number of clicks for each ip address, then obtains the maximum of that.
18		NoOfIPs	For a given partner, count the number of unique ip address.

19		IP/ClickRatio	For a given partner, count the number unique ip address and take divide by no of click from that partner.
20		IPVar	For a given partner, count the number of clicks for each ip address, then obtains the Variance of that.
21		IPEntropy	For a given partner, count the number of clicks for each ip address, then obtains the Entropy of that.
22	**Agent**	MaxSameAgentCount	For a given partner, count the number of clicks for each Agent, then obtains the maximum of that.
23		NoOfAgents	For a given partner, count the number of unique Agents.
24		Click/AgentRatio	For a given partner, count the number unique Agent and take divide by no of click from that partner.
25		AgentVar	For a given partner, count the number of clicks for each Agent, then obtains the Variance of that.
26		AgentEntropy	For a given partner, count the number of clicks for each Agent, then obtains the Entropy of that.
27	**Campaign ID**	MaxSameCidCount	For a given partner, count the number of clicks for each Campaign ID, then obtains the maximum of that.
28		NoOfCid	For a given partner, count the number of unique Campaign ID.
29		Cid/ClickRatio	For a given partner, count the number unique Campaign ID and take divide by no of click from that partner.
30		CidVar	For a given partner, count the number of clicks for each Campaign ID, then obtains the Variance of that.
31		CidEntropy	For a given partner, count the number of clicks for each Campaign ID, then obtains the Entropy of that.
32	**Country**	MaxSameCntrCount	For a given partner, count the number of clicks for each Country, then obtains the maximum of that.
33		NoOfCountries	For a given partner, count the number of unique Country.
34		Country/ClickRatio	For a given partner, count the number unique Country and take divide by no of click from that partner.
35		CntrVar	For a given partner, count the number of clicks for each Country, then obtains the Variance of that.
36		CountryEntropy	For a given partner, count the number of clicks for each Country, then obtains the Entropy of that.
37	**Partner**	TotalClicks	For a given partner, count total no of clicks.
38	**Referrer**	NoOfRefers	For a given partner, count the number of unique referrer.
39		Refer/ClickRatio	For a given partner, count the number of unique referrer and divide by no of click for that partner.
40		NonreferedClickRatio	For a given partner, count the number of non-referred click and divide by no of click for that partner.
41	**Category**	CategoryPrior	Probability of being fraud if the click is for that category. This probability is calculated from training set and used for validation and test set.