**Data Challenge: Fraud Detection**

**Task:**

E-commerce websites often transact huge amounts of money. And whenever a huge amount of money is moved, there is a high risk of users performing fraudulent activities, e.g. using stolen credit cards, doing money laundry, etc. Machine Learning really excels at identifying fraudulent activities. Any website where you put your credit card information has a risk team in charge of avoiding frauds via machine learning.

**The goal of this challenge is to build a simple machine learning model that predicts the probability that the first transaction of a new user is fraudulent.** Feel free to use any programming language you’re mostly comfortable with to tackle the problem.

**Data:**

**Fraud\_Data.zip file** contains information about each user first transaction.

**Columns**:  
**user\_id** : Id of the user. Unique by user  
**signup\_time** : the time when the user created her account (GMT time)  
**purchase\_time** : the time when the user bought the item (GMT time)  
**purchase\_value** : the cost of the item purchased (USD)  
**device\_id** : the device id. You can assume that it is unique by device. I.e., 2 transactions  
with the same device ID means that the same physical device was used to buy  
**source** : user marketing channel: ads, SEO, Direct (i.e. came to the site by directly typing  
the site address on the browser).  
**browser** : the browser used by the user.  
**sex** : user sex: Male/Female  
**age** : user age  
**ip\_address** : user numeric ip address  
**class** : this is what we are trying to predict: whether the activity was fraudulent (1) or not  
(0).