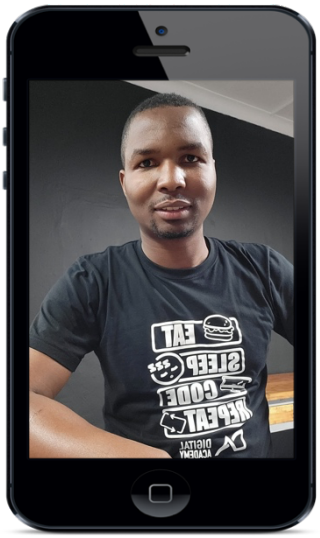
DIGITAL ACADEMY

SABELO MSIMANGO



During the 6 month period of the IT: Systems Support learnership programme at The Digital Academy **Sabelo** and his team created a mobile application called **E-commerce Fraud Predictions.**

Sabelo's main responsibilities on the project were Model building, Model deployment, and Data Visualisation.

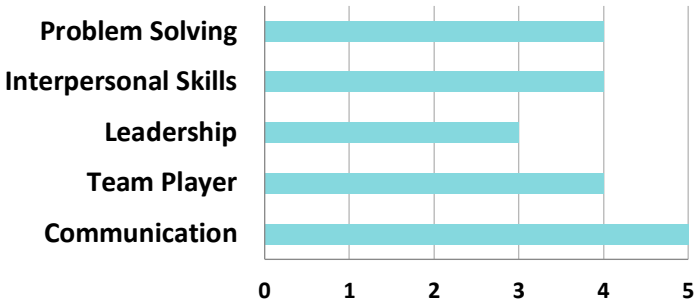
I would have no hesitation in recommending **Sabelo** as Junior Data Scientist. He has demonstrated strong technical skills and understanding of the technologies he has used. He is able to work individually and in a team with equal proficiency.

SOFT SKILLS

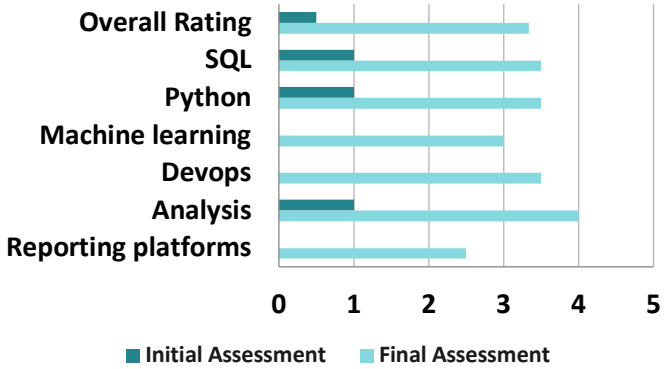
Sabelo's most improved competency:

Team player

The skill to venture on with strong resolve and persistence, committed to getting the job done and works well within a team.



TECHNICAL SKILLS



Primary Skills

SQL
Python
Analysis
DevOps

Secondary Skills

Machine Learning
Reporting platforms

Dylan Knevitt
(Technical Manager)

Melisha Moodley
(Programme Manager)

Gary Bannatyne
(MD & Co-founder)

061 544 9366

info@thedigitalacademy.co.za

@StreetDev

Eccomerce Fraud Prediction

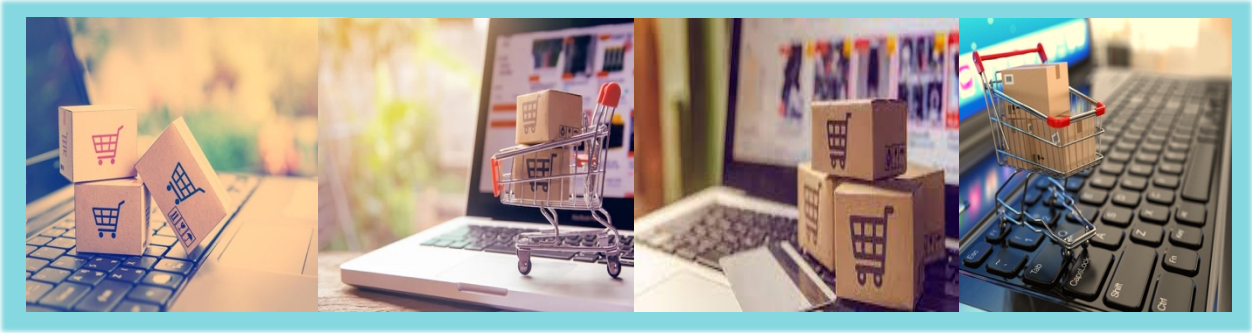
Promo pack

Problem Statement

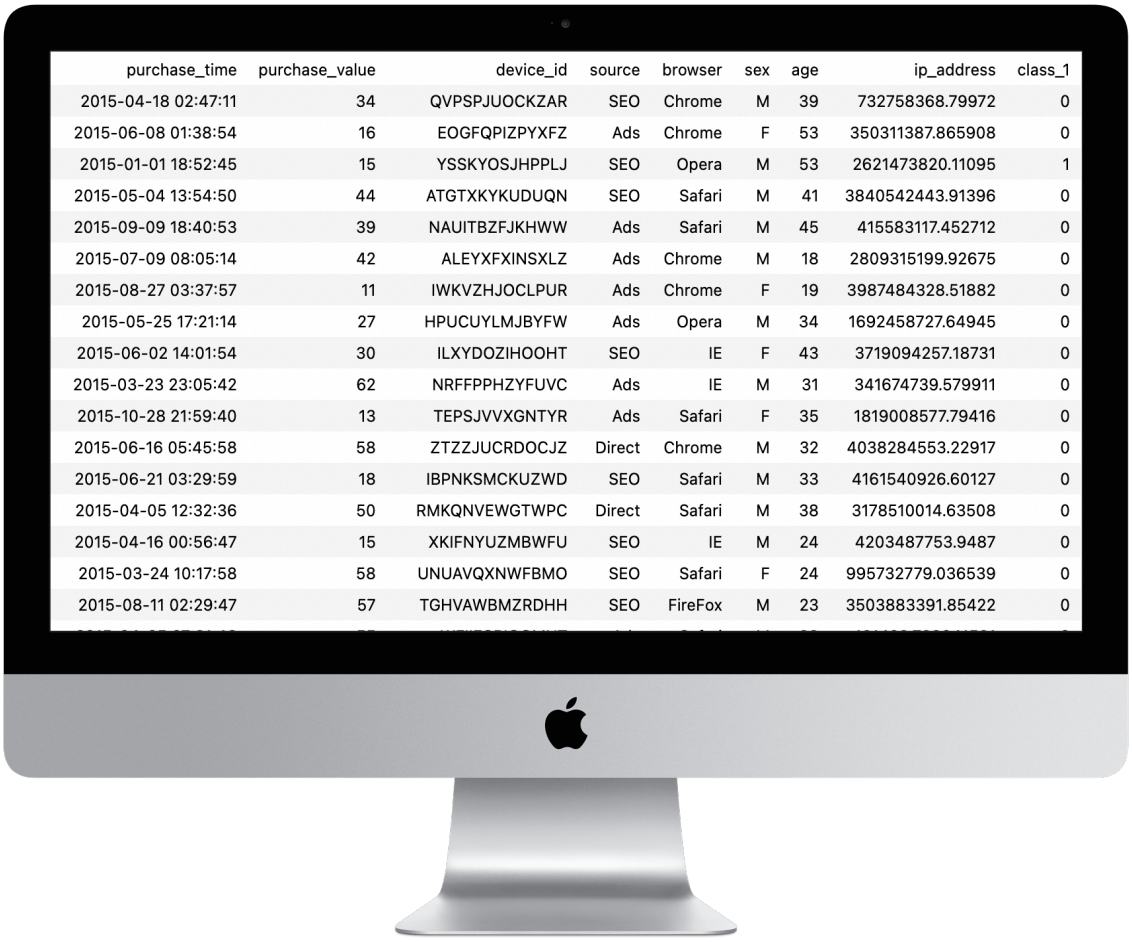
For fraud detection on e-commerce platforms, many fraud detection algorithms have focused on financial crimes such as credit card frauds and risky accounts. Online fraud occurs when fraudsters use digital channels to infiltrate a customers' online account to make fraudulent purchases.

Solution

To develop a model that identifies fraudulent activities in a timely manner to prevent any theft.



Data Set



purchase_time	purchase_value	device_id	source	browser	sex	age	ip_address	class_1
2015-04-18 02:47:11	34	QVPSPJUOCKZAR	SEO	Chrome	M	39	732758368.79972	0
2015-06-08 01:38:54	16	EOGFQPIZPYXFZ	Ads	Chrome	F	53	350311387.865908	0
2015-01-01 18:52:45	15	YSSKYOSJHPPLJ	SEO	Opera	M	53	2621473820.11095	1
2015-05-04 13:54:50	44	ATGTXKYKUDUQN	SEO	Safari	M	41	3840542443.91396	0
2015-09-09 18:40:53	39	NAUITBZFJKHWW	Ads	Safari	M	45	415583117.452712	0
2015-07-09 08:05:14	42	ALEYXFXINSXLZ	Ads	Chrome	M	18	2809315199.92675	0
2015-08-27 03:37:57	11	IWKVZHJOCLPUR	Ads	Chrome	F	19	3987484328.51882	0
2015-05-25 17:21:14	27	HPUCUYLMJBYFW	Ads	Opera	M	34	1692458727.64945	0
2015-06-02 14:01:54	30	ILXYDOZIHOOHT	SEO	IE	F	43	3719094257.18731	0
2015-03-23 23:05:42	62	NRFFPPHZYFUVC	Ads	IE	M	31	341674739.579911	0
2015-10-28 21:59:40	13	TEPSJVVXGNTYR	Ads	Safari	F	35	1819008577.79416	0
2015-06-16 05:45:58	58	ZTZZJUCRDOCJZ	Direct	Chrome	M	32	4038284553.22917	0
2015-06-21 03:29:59	18	IBPNKSMCKUZWD	SEO	Safari	M	33	4161540926.60127	0
2015-04-05 12:32:36	50	RMKQNVEWGTWPC	Direct	Safari	M	38	3178510014.63508	0
2015-04-16 00:56:47	15	XKIFNYUZMBWFU	SEO	IE	M	24	4203487753.9487	0
2015-03-24 10:17:58	58	UNUAVQXNWFBMO	SEO	Safari	F	24	995732779.036539	0
2015-08-11 02:29:47	57	TGHVAWBMZRDHH	SEO	FireFox	M	23	3503883391.85422	0

Fraud data containing a sample of users with information such as user id, gender, age, signup_time, purchase time, device id, ip_address, and whether they committed fraud or not. it has 151112 rows and 11 columns.

Analysis Dashboard

We compare the total number of IP addresses per browser to check which browser was used the most and which one was used less, it also counts the total number of IP addresses that were used per browser. We also compared the number of purchased items in each browser, and the source of transactions that help us to determine whether you use an Ad which is a pop-up on the internet, SEO straight to search engine, and direct which is the actual application. Classification of fraud is the actual feature that determines if a person has committed fraud or not.

Model Predictions

A User is requested to enter a number of features in the UI. These features will then be fed into a machine learning model trained with a historical Dataset in order to classify a transaction as fraudulent or not.