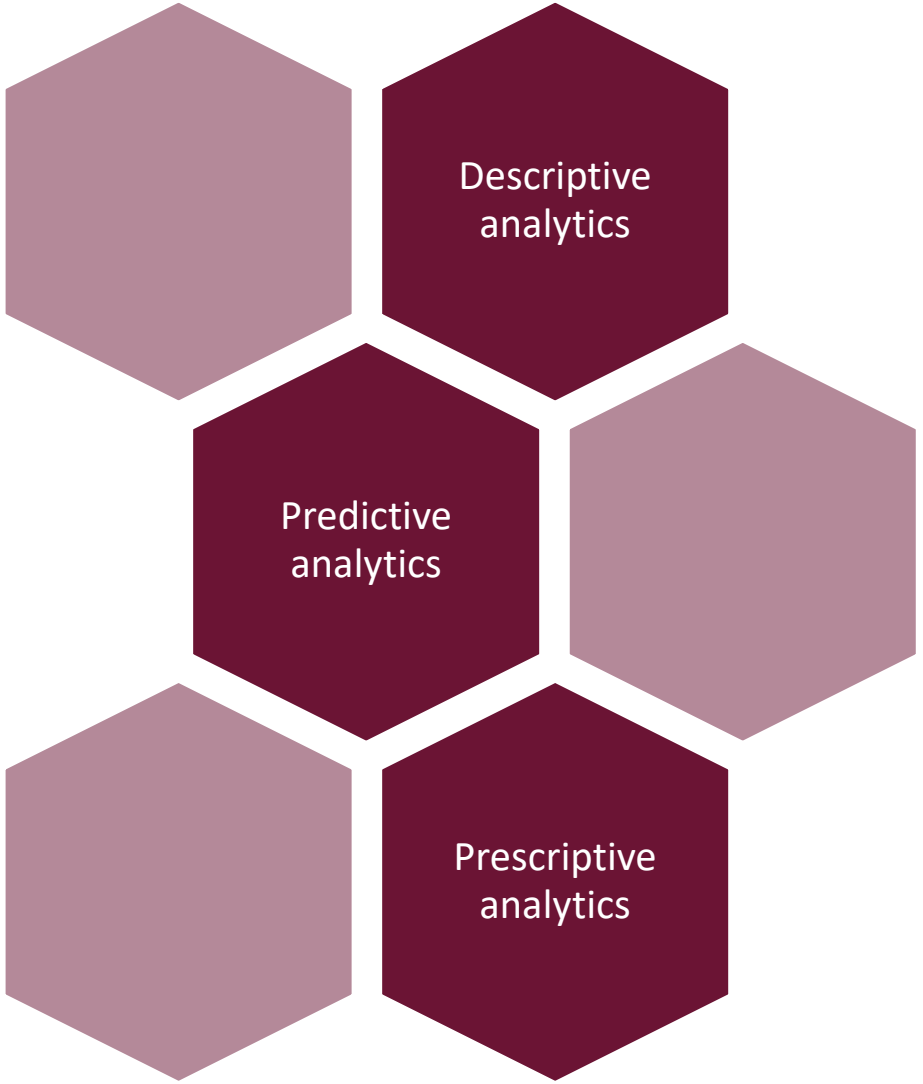
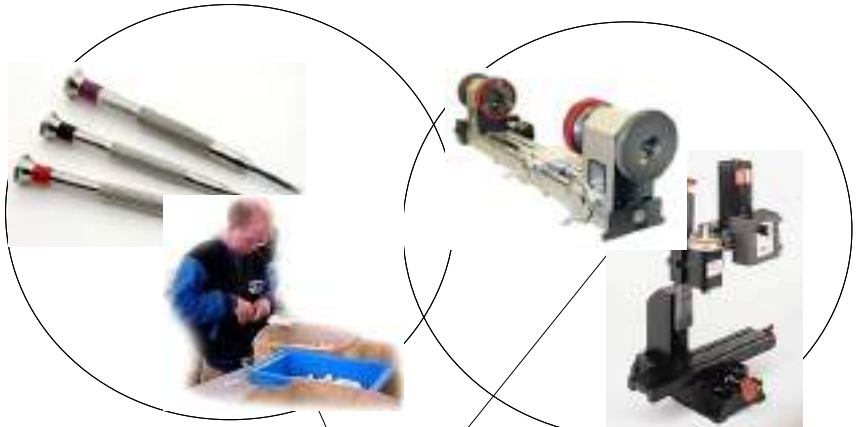
Supply chain data science: Unleashing AI in the business domain

Jacomine Grobler
Department of Industrial Engineering
jacominegrobler@sun.ac.za

Data science

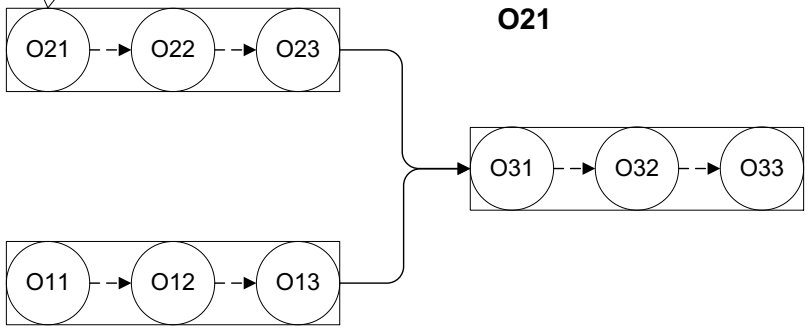


Prescriptive analytics projects



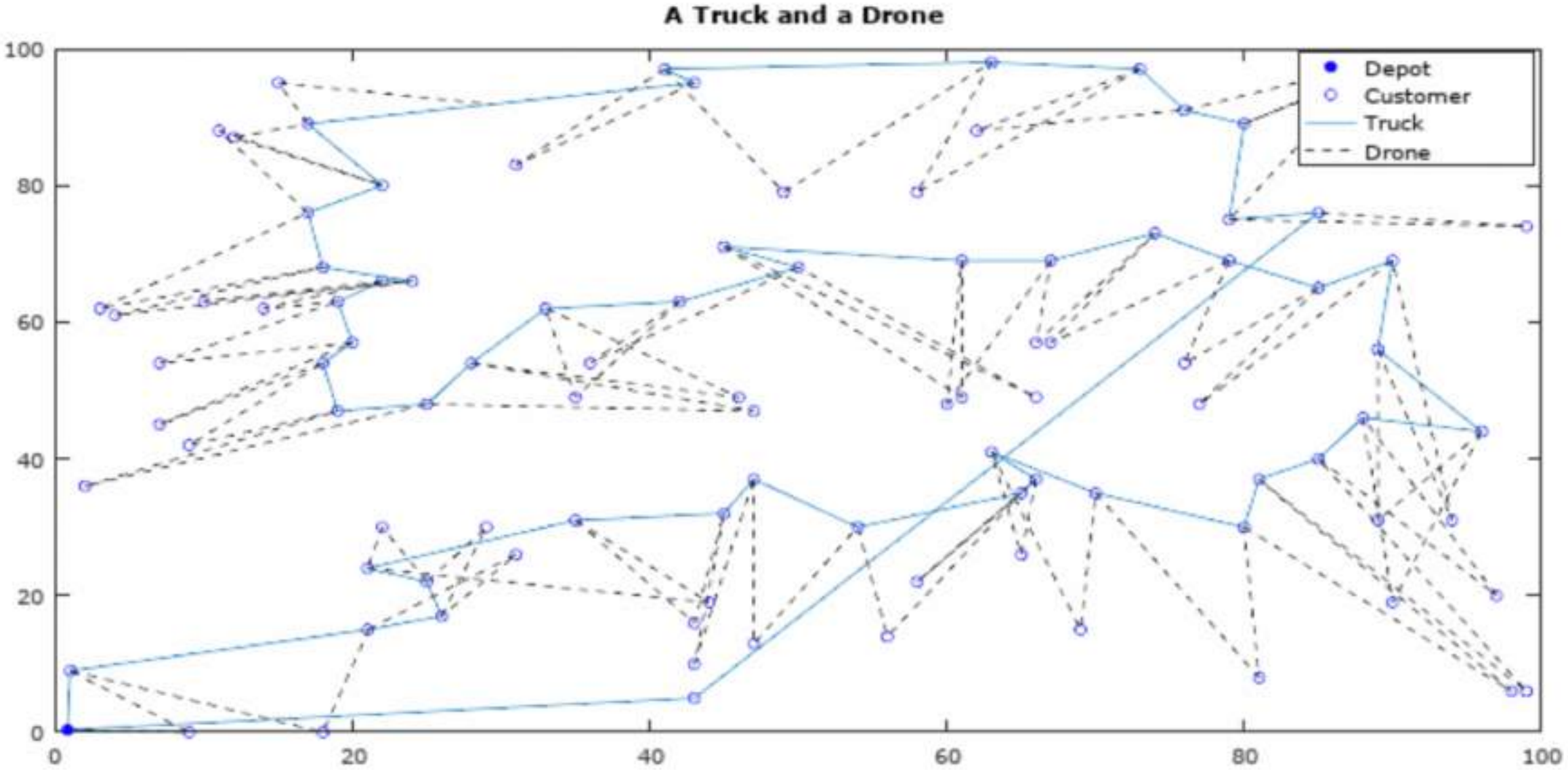
Auxilliary Resource Set for O21

Primary Resource Set for O21



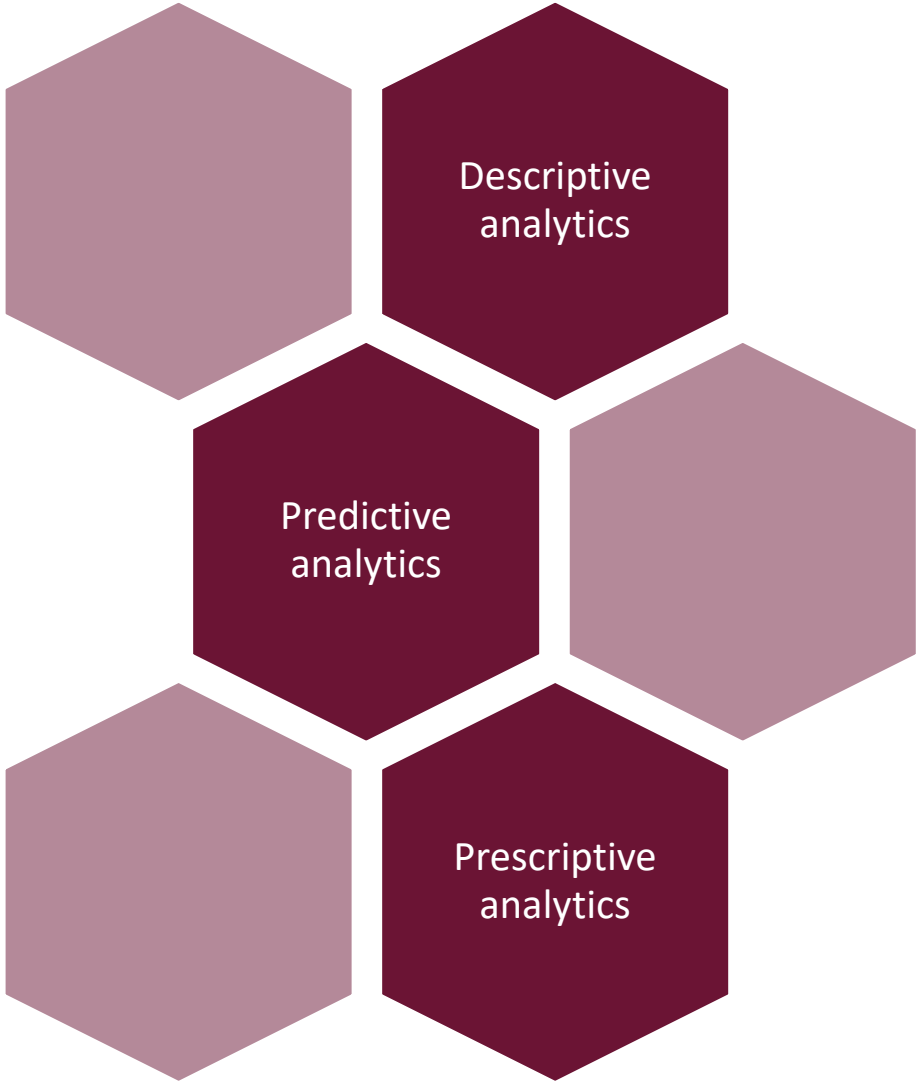
Minimize
- Makespan
- Queue time
Maximize
-Customer service

Preliminary Results



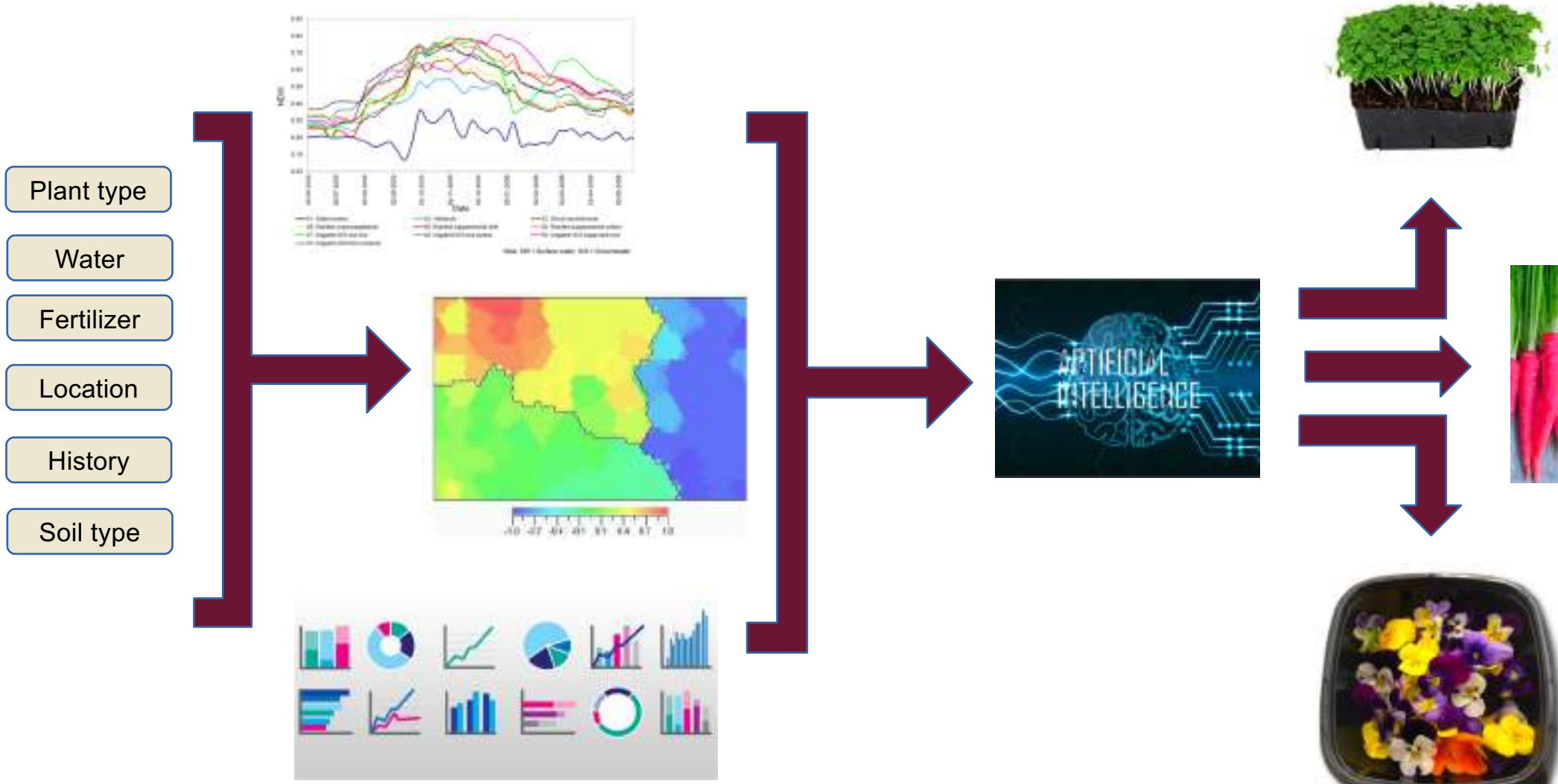
Acknowledgements: Tsietsi Moremi

Data Science



Acknowledgements: Yolandi Le Roux; University of Pretoria

A Predictive Model from the Agricultural Industry



Acknowledgements: Yolandi Le Roux; University of Pretoria

Acknowledgements: Yolandi Le Roux; University of Pretoria

Classifier Type	Algorithm	Accuracy
Bayesian	Bayes Net	40.89%
	Naïve Bayes	36.89%
	Naïve Bayes Updateable	36.89%
Functions	Logistic	35.15%
	Multi layer Perception	55.92%
	Simple Logistic	33.79%
	SMO	38.34%
Lazy	IBK	93.90%
	Star	90.16%
	LWL	30.51%
Meta	Attribute Selected Classifier	81.79%
	Bagging	73.59%
	Classification Via Regression	74.59%
	Filtered Classifier	61.11%
	Iterative Classifier Optimiser	44.72%
	Logit Boost	44.72%
	Multi Class Classifier	34.15%
	Random Committee	94.17%
	Randomizable Filtered Classifier	94.17%
	Random Sub Space	68.49%
Rules	Decision Table	63.48%
	JRip	68.76%
	Part	88.25%
Trees	Hoeffding Tree	36.16%
	J48	87.70%
	LMT	88.98%
	Random Forest	94.17%
	Random Tree	93.90%
	REP Tree	63.39%

Results

UNIVERSITEIT iYUNIVESITHI STELLENBOSCH UNIVERSITY

94% accuracy obtained with a random forest algorithm

Acknowledgements: Yolandi Le Roux; University of Pretoria

A job shop scheduling problem with due dates: Which rule should we use?

Operation 5

Operation 6

Operation 7

- FIFO
- EDD
- SPT

Resources	Time (Days)													
	12	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*
1	Operation 1				Operation 3			Operation 4						
2	Operation 2													
3								Operation 4						
4	Operation 1													

Acknowledgements: M Agigi; University of Pretoria

The training data

- Five data sets of different sizes
  - 56, 100, 146, 200, 256-operations
- Three algorithms
  - FIFO, SPT & EDD

At each “scheduling decision” the Work in Process (WIP) and Average Remaining Processing time (ARP) was calculated and the best performing algorithm (wrt makespan) was recorded

OPS	WIP	ARP	Best algorithm
56	0	9.79	FIFO
56	23.6	8.76	SPT
100	6426.52	98.59	EDD
100	6333.05	98.59	EDD
146	14719.62	77.53	FIFO
146	0	98.9	FIFO
200	358.18	10.79	FIFO
200	439.07	10.42	EDD
256	1454	33.93	SPT

Acknowledgements: M Agigi; University of Pretoria

Results

Classifier Type	Algorithm	Accuracy (SD)	Training Time (ms)	Testing Time (ms)	Total Time (ms)
Bayesian	Naive Bayes	61.74 (2.26)	2.08	0.94	3.02
	Bayes Net	87.59 (2.35)	5.00	0.63	5.63
	Naive Bayes Multinomial	37.95 (3.47)	0.31	0.10	0.42
Functions	Logistic	57.21 (2.19)	44.27	0.16	44.43
	Multilayer Perceptron	73.95 (3.34)	1012.66	0.57	1013.23
	SMO	60.08 (2.49)	103.80	0.26	104.06
	Simple Logistic	60.24 (2.30)	256.72	0.21	256.93
Lazy	IB1	93.87 (1.60)	0.16	21.72	21.88
	IBk	94.26 (1.44)	0.26	20.63	20.89
	KStar	90.30 (1.90)	0.05	804.27	804.32
	LWL	65.60 (1.88)	0.10	389.38	389.48
Rules	DecisionTable	90.82 (1.96)	19.32	0.31	19.64
	JRip	90.68 (1.91)	232.92	0.16	233.07
	MODLEM	84.32 (2.26)	121.61	3.13	124.74
	OneR	84.55 (2.16)	1.77	0.16	1.93
	PART	92.58 (1.85)	28.18	0.36	28.54
	ZeroR	53.29 (0.19)	0.05	0.05	0.10
Trees	DecisionStump	61.14 (2.29)	0.57	0.10	0.68
	HoeffdingTree	61.77 (2.26)	6.61	0.63	7.24
	J48	93.36 (1.54)	13.33	0.10	13.44
	LMT	93.16 (1.57)	453.80	0.47	454.27
	RandomForest	94.13 (1.49)	215.63	11.77	227.40
	RandomTree	93.95 (1.56)	4.22	0.31	4.53
	REPTree	91.59 (1.78)	4.32	0.21	4.53
	SimpleCart	93.16 (1.67)	84.87	0.10	84.97

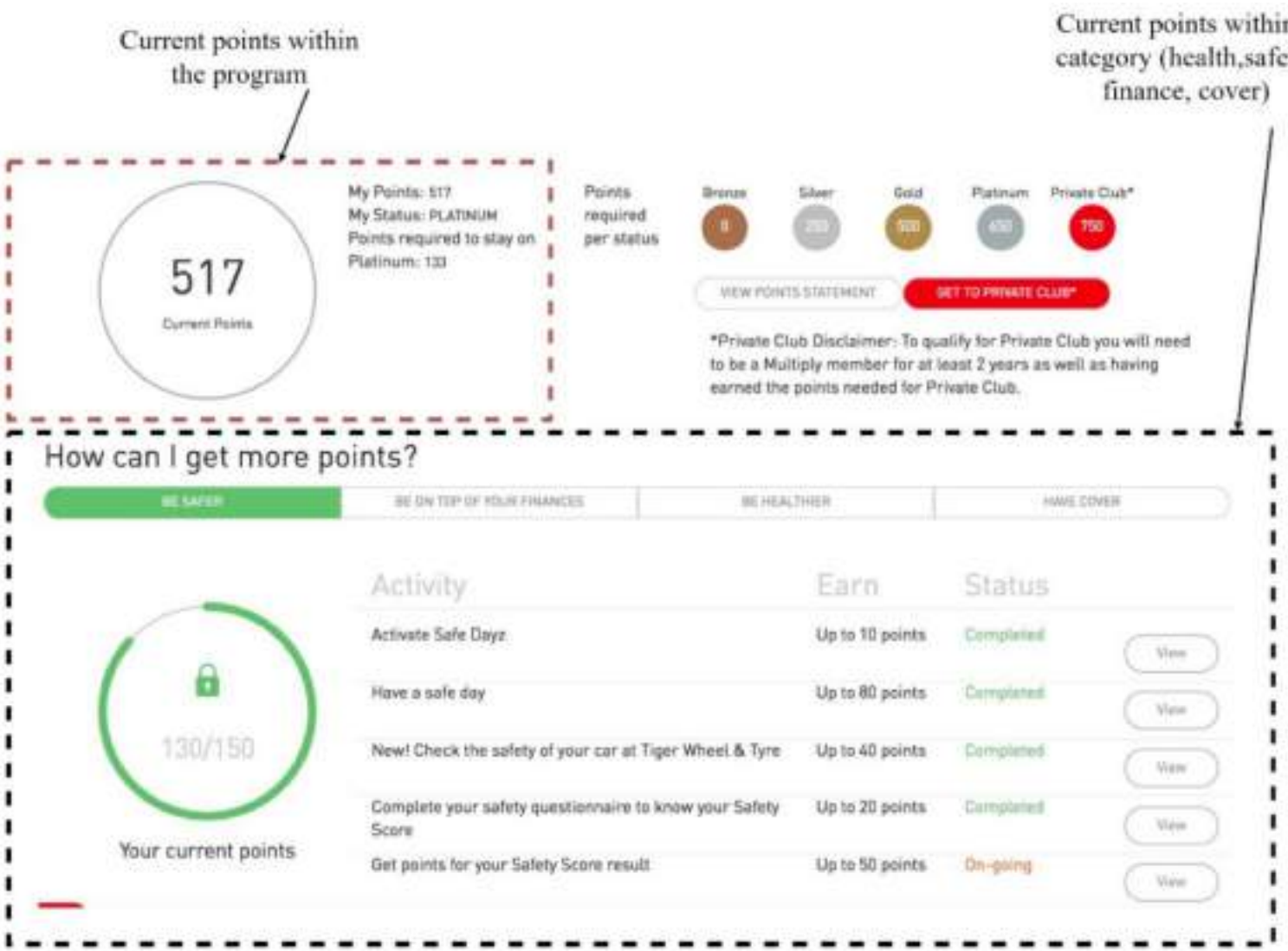
Acknowledgements: M Agigi; University of Pretoria

Results

- Based on problem size, APT & WIP, the correct rule could be selected with an accuracy of **94%**.

- Future work can include incorporating more attributes, utilizing a larger dataset and investigating more complex scheduling algorithms.

Collaborative Filter Type Recommender for Incentive Programmes



Classify customers into incentive categories

↓

KNN clustering to identify similar users

↓

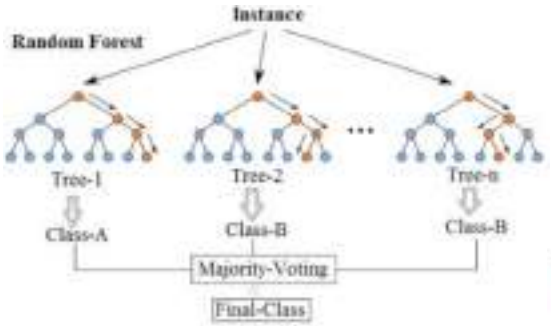
Identify recommendations based on all point contributing activities

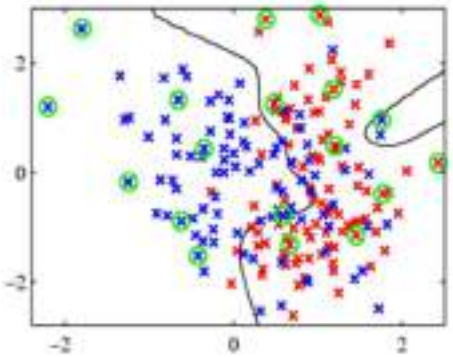
Acknowledgements: Ridhaa Beneveld; University of Pretoria

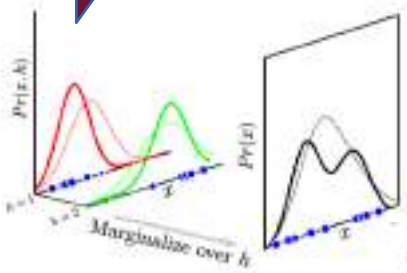
69% of users ascended in the programme

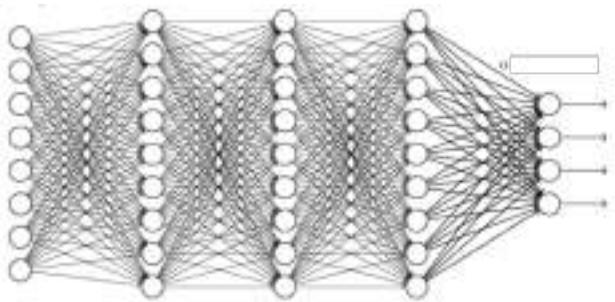
Customer Segmentation by means of Data Science

- Order qty
- Location
- Credit history
- $ contribution
- Payment type
- Products
- Complaints
- Socio-demographics
- Industry
- Cost 2 serve





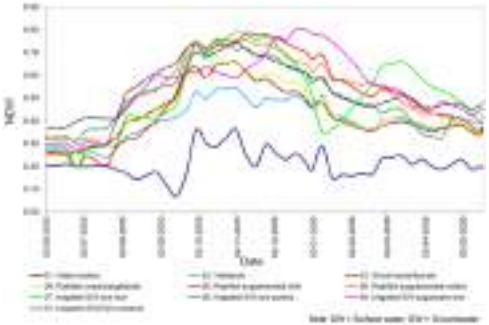


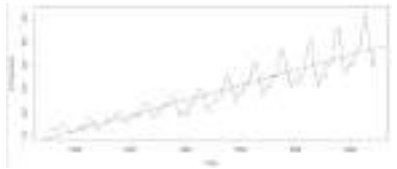


- Unique product service agreements
- Business rules
- Resource planning

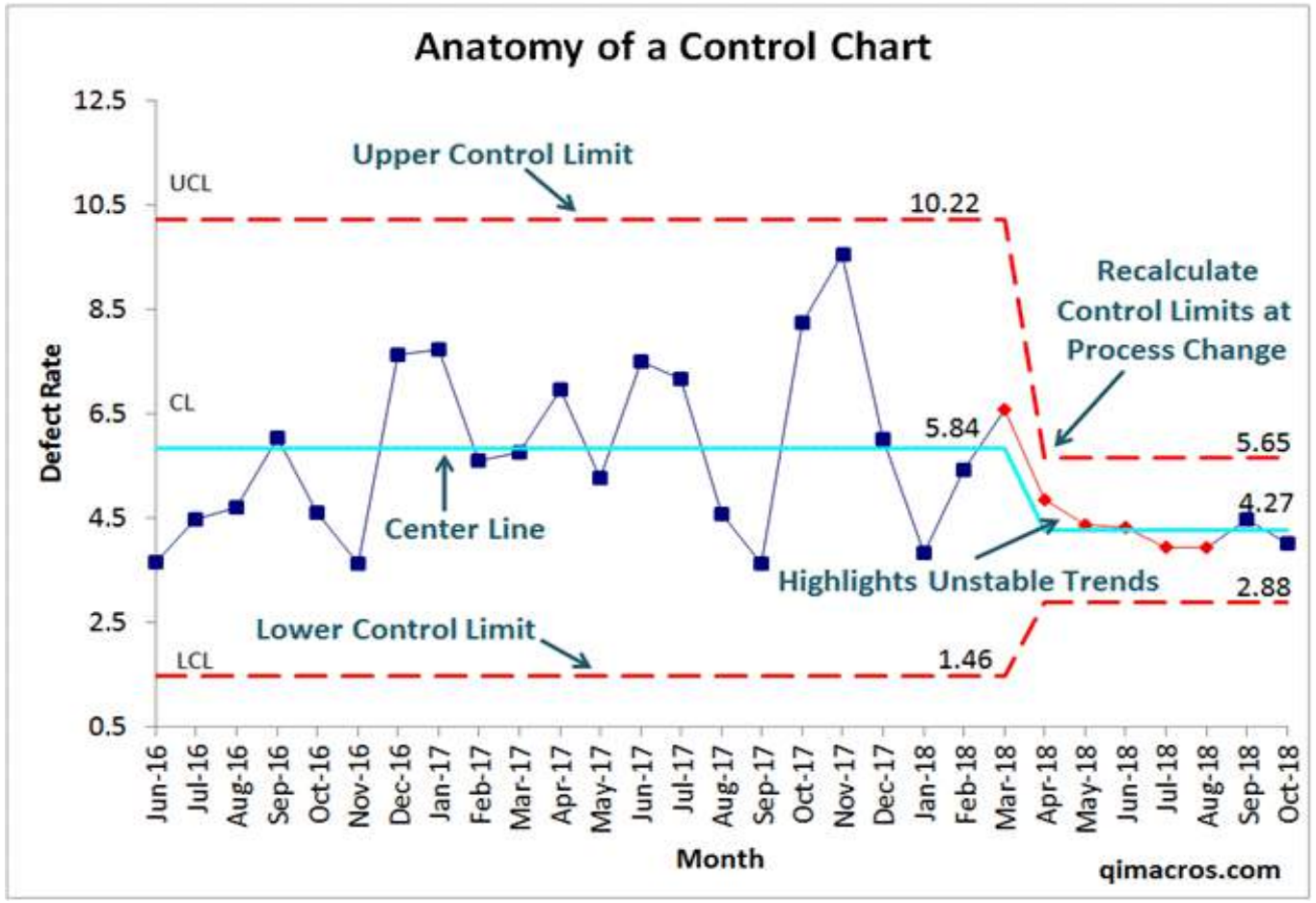
Predicting Delivery Times

- Location
- Driver
- Traffic
- Weather
- Residence type
- Order $
- Socio-demographics
- Distance





Predicting Manufacturing Performance



Clustering of production processes

↓

Classification of processes wrt quality

↓

Training & comparison with SPC charts

Acknowledgements: Sibusiso Khoza

Other Projects

- Predicting port delays from wind, wave and other data
- Predicting energy requirements in the hospitality industry
- Greenhouse gas prediction
- Supply chain performance prediction
- Predicting diabetes through medication purchasing data

Acknowledgements: Andries Engelbrecht; Cecil Musisinyani; Philip du Plooy; Lumi Dreyer; University of Pretoria

Supply chain data science: Unleashing AI in the business domain

Jacomine Grobler
Department of Industrial Engineering
jacominegrobler@sun.ac.za