

## Johnson Osioma MSc in Big Data

### Problem Definition

A **Max SAT** problem is an optimised version of SAT problem. The input to Max SAT problem propositional logic formulae in Conjunctive Normal Form: this is a conjunction of clauses where each clause is a disjunction of values which are variables or their negations.

### Project Description

This project considers existing datasets of features describing instances of optimisation problems paired with the performance of a set of algorithms to solve them. The project aim is to produce both (i.) visual maps of algorithms to solve them (ii.) predictive models of algorithm performance.

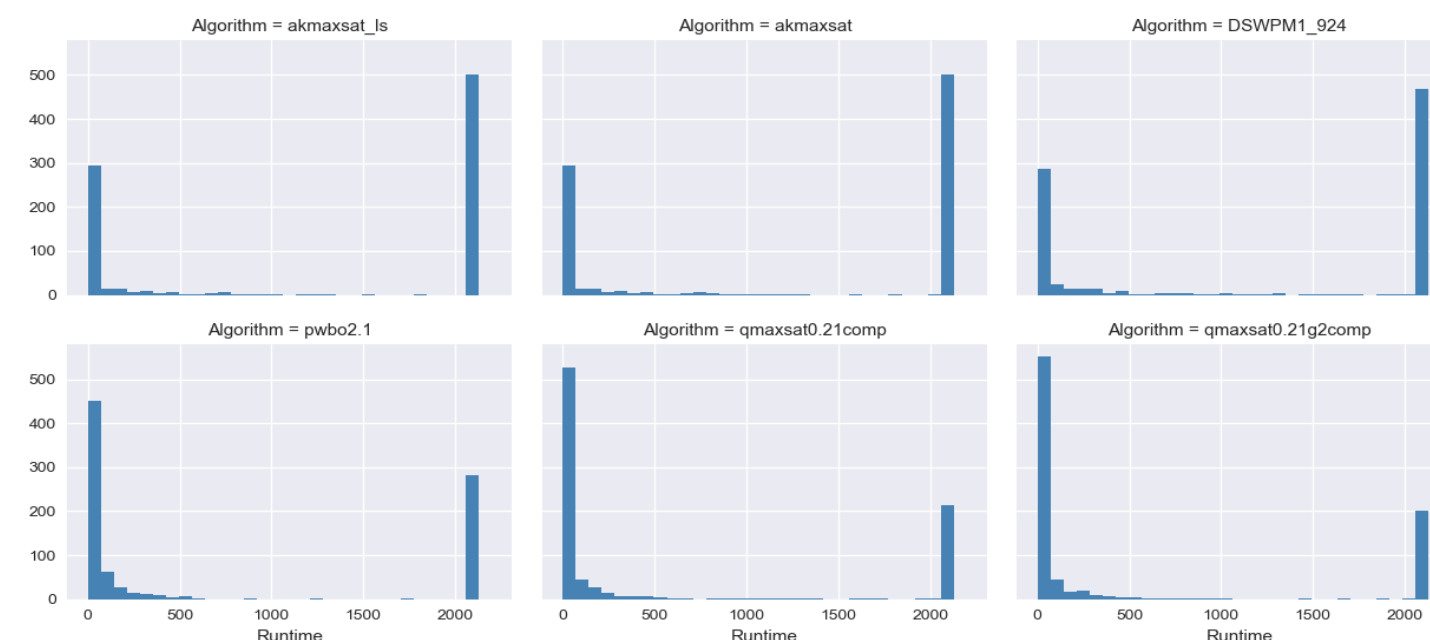
The performance data is based on the algorithms: akmaxsat, akmaxsat\_ls, DSWPM1\_924, pwbo2.1, qmaxsat0.21comp, qmaxsat0.21g2comp.

The approach is to use libraries in Python for visualization, and prediction using machine learning models. This is accomplished by applying a number of methods and assess how they tackle this problem.

### Discussion

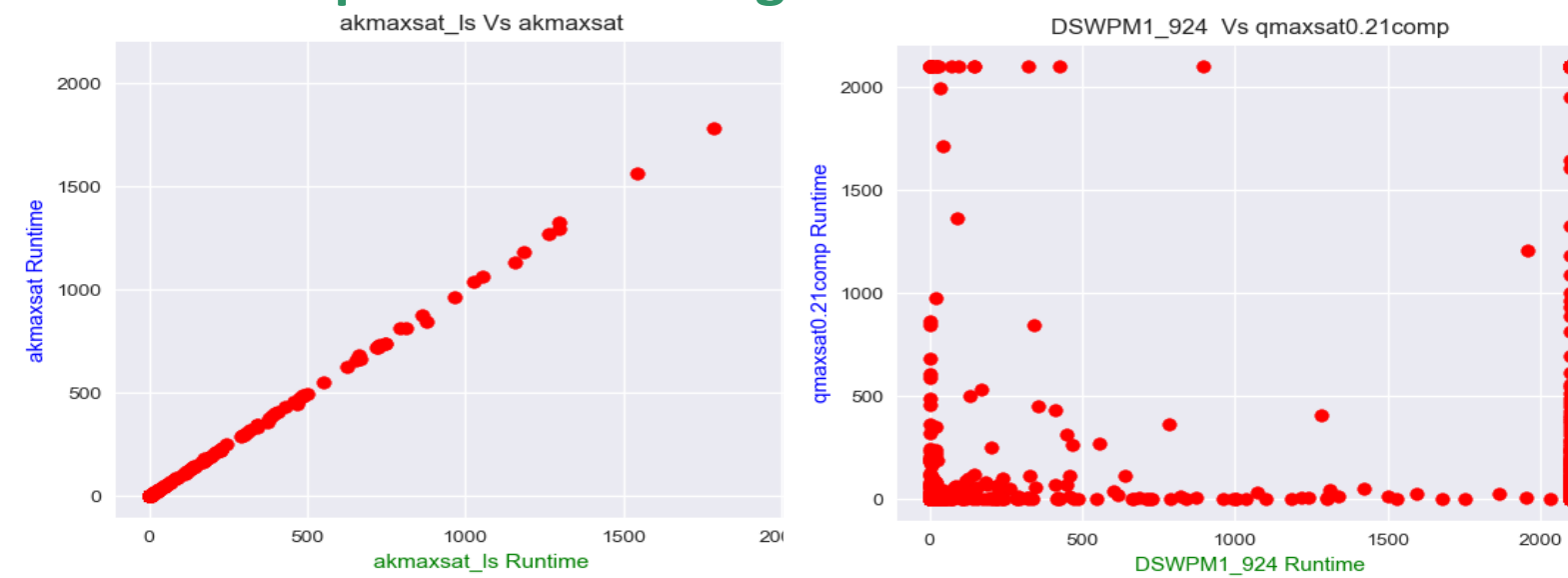
Further work on the project includes visualisation with t-NSE plots for non-linear high dimensionality reduction and algorithm predication

### Runtime performance of algorithms



Here you see similarity in runtime for the first three algorithms runtimes

### Scatter plots of some algorithms

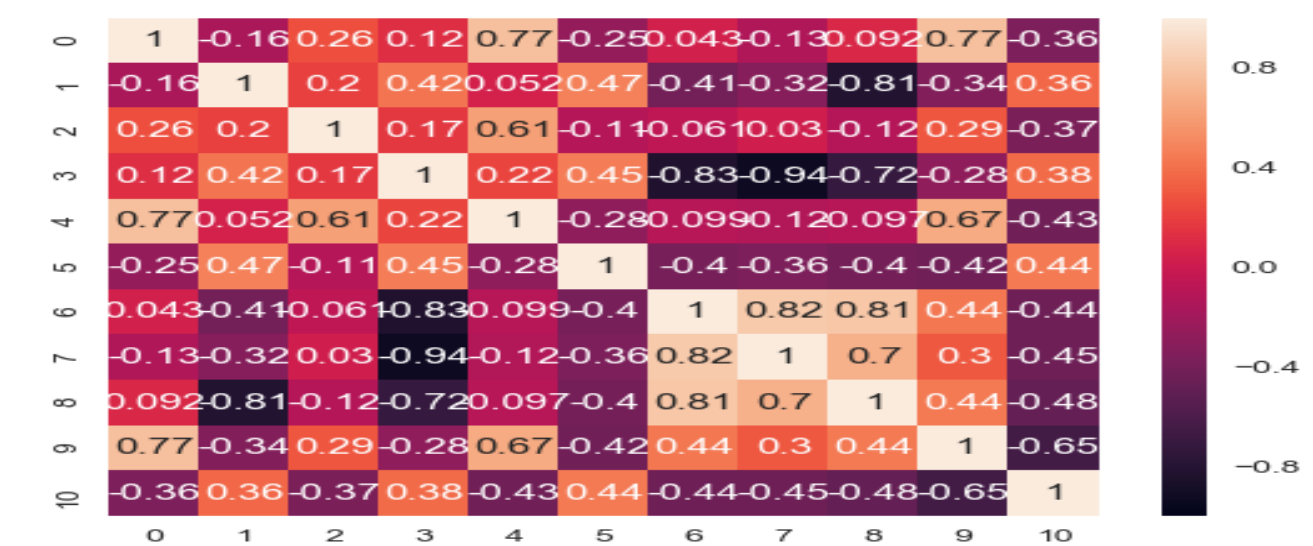


Whilst there is a great similarity between akmaxsat and akmaxsat\_ls runtime, there is a wide disparity between DSWPMI\_924 and qmaxsat0.21comp runtime- this shows that whilst some instances timed out with qmaxsat0.21comp, they actually did very well with DSWPMI\_924 and vice versa.

### Table analysis of best algorithm

Rank	Weight	Wghtd Score (akmaxsat)		Wghtd Score (akmaxsat_ls)		Wghtd Score (DSWPM1_924)		Wghtd Score (pwbo2.1)		Wghtd Score (qmaxsat0.21comp)		Wghtd Score (qmaxsat0.21g2comp)		Wghtd Score (p)							
		(%)		(%)		(%)		(%)		(%)		(%)									
1	6	285	32.53	1710	46	5.25	276	78	8.90	468	21	2.40	126	192	21.92	1152	254	29.00	1524		
2	5	99	10.62	465	287	32.76	1495	23	2.63	115	63	7.19	315	212	24.20	1060	198	22.60	990		
3	4	65	7.42	260	61	6.96	244	257	29.34	1028	282	32.19	1128	157	17.92	628	54	6.16	216		
4	3	171	19.52	513	76	8.68	228	180	20.55	540	271	30.94	813	67	7.65	201	111	12.67	333		
5	2	244	27.85	488	179	20.43	358	57	6.51	114	151	17.24	302	204	23.29	408	41	4.68	82		
6	1	18	2.05	18	227	25.91	227	281	32.08	281	88	10.05	88	44	5.02	44	218	24.89	218		
		876	100.00	3454	876	100.00	2768	876	100.00	2546	876	100.00	2772	876	100.00	3493	876	100.00	3363		
				2nd			4th						5th			6th			1st		3rd

### Top 10 correlations



Here you observe that the correlations get stronger as the legend grows darker with -0.65 being the strongest correlation runtime value

### Analysis of feature values using correlation

	akmaxsat	akmaxsat_ls	DSWPM1_924	pwbo2.1	qmaxsat0.21comp	qmaxsat0.21g2comp
Feature 1	vsg_cls_min (-0.65)	vsg_cls_min (-0.65)	binary (0.43)	perc_soft (0.26)	perc_soft (0.36)	perc_soft (0.43)
Feature 2	pnr_var_min (-0.48)	pnr_var_min (-0.48)	horn_max (0.38)	binary (0.22)	unary (0.21)	pnr_var_mean (0.34)
Feature 3	pnr_cls_mean (-0.45)	pnr_cls_mean (-0.45)	pnr_var_spread (-0.37)	horn (-0.21)	pnr_var_mean (0.2)	unary (0.28)
Feature 4	pnr_var_mean (-0.44)	pnr_var_mean (-0.45)	horn (-0.37)	numClauses (0.2)	binary (0.19)	horn (-0.27)
Feature 5	vsg_var_std (0.44)	vsg_var_std (0.44)	horn_spread (-0.36)	horn_max (-0.2)	horn (-0.17)	pnr_var_max (0.26)
Feature 6	horn_mean (-0.43)	horn_mean (-0.43)	var_clauses_ratio (-0.34)	horn_spread (-0.19)	horn_mean (-0.17)	pnr_cls_mean (0.19)
Feature 7	horn (0.38)	horn (0.38)	horn_mean (-0.32)	pnr_var_mean (0.19)	numClauses (0.17)	pnr_var_min (0.17)
Feature 8	horn_std (-0.37)	horn_std (-0.37)	pnr_var_min (0.31)	pnr_cls_std (0.16)	horn_max (-0.17)	binary (0.17)
Feature 9	pnr_var_spread (0.36)	pnr_var_spread (0.36)	horn_std (-0.3)	horn_mean (-0.15)	pnr_cls_std (0.16)	horn_max (-0.13)
Feature 10	horn_min (-0.36)	horn_min (-0.36)	pnr_var_std (-0.28)	horn_min (-0.13)	horn_spread (-0.16)	horn_spread (-0.12)

Here you observe that Akmaxsat and akmaxsat\_ls have almost identical top 10 features and values. DSWPM1\_924 has a feature listed (pnr\_var\_spread), which has negative correlation where akmaxsat and akmaxsat\_ls show positive correlation values for the same feature. This is an interesting finding that needs further research.