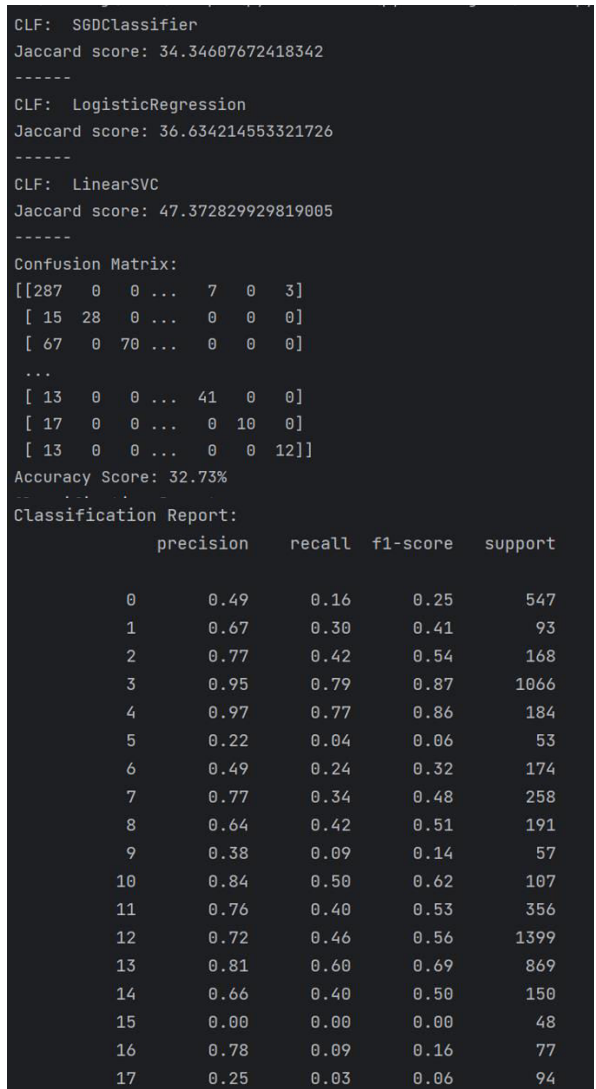


Project Development Phase Model Performance Test

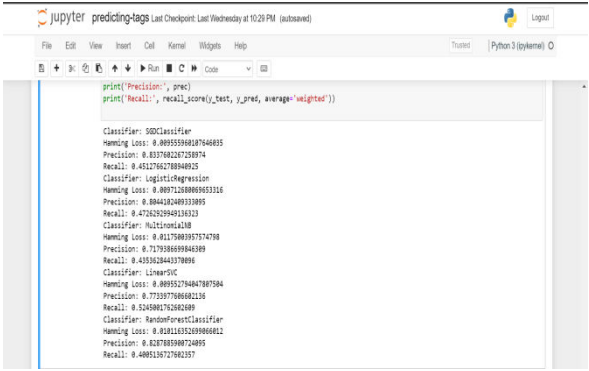
Date	19 May 2023
Team ID	NM2023TMID06421
Project Name	Autonomous Tagging Of Stack Overflow Questions

Model Performance Testing:

S.No.	Parameter	Values	Screenshot
1.	Metrics	Classification Model: Confusion Matrix Accuray Score Classification Report	 <pre> CLF: SGDClassifier Jaccard score: 34.34607672418342 ----- CLF: LogisticRegression Jaccard score: 36.634214553321726 ----- CLF: LinearSVC Jaccard score: 47.372829929819005 ----- Confusion Matrix: [[287 0 0 ... 7 0 3] [15 28 0 ... 0 0 0] [67 0 70 ... 0 0 0] ... [13 0 0 ... 41 0 0] [17 0 0 ... 0 10 0] [13 0 0 ... 0 0 12]] Accuracy Score: 32.73% ----- Classification Report: precision recall f1-score support 0 0.49 0.16 0.25 547 1 0.67 0.30 0.41 93 2 0.77 0.42 0.54 168 3 0.95 0.79 0.87 1066 4 0.97 0.77 0.86 184 5 0.22 0.04 0.06 53 6 0.49 0.24 0.32 174 7 0.77 0.34 0.48 258 8 0.64 0.42 0.51 191 9 0.38 0.09 0.14 57 10 0.84 0.50 0.62 107 11 0.76 0.40 0.53 356 12 0.72 0.46 0.56 1399 13 0.81 0.60 0.69 869 14 0.66 0.40 0.50 150 15 0.00 0.00 0.00 48 16 0.78 0.09 0.16 77 17 0.25 0.03 0.06 94 </pre>

			17	0.25	0.03	0.06	94
			18	0.75	0.63	0.68	374
			19	0.67	0.27	0.38	60
			20	0.37	0.11	0.17	102
			21	0.52	0.24	0.33	46
			22	0.67	0.29	0.40	84
			23	0.96	0.54	0.70	90
			24	0.78	0.14	0.24	50
			25	0.93	0.73	0.82	162
			26	0.84	0.55	0.66	153
			27	0.89	0.57	0.70	42
			28	0.85	0.49	0.62	67
			29	0.62	0.23	0.33	66
			30	0.79	0.50	0.61	54
			31	0.50	0.02	0.04	47
			32	0.68	0.23	0.35	56
			33	0.70	0.34	0.46	67
			34	0.94	0.84	0.89	201
			35	0.75	0.50	0.60	54
			36	0.69	0.30	0.42	67
			37	0.92	0.62	0.74	132
			38	0.95	0.68	0.79	59
			38	0.95	0.68	0.79	59
			39	0.50	0.26	0.35	423
			40	0.74	0.29	0.42	121
			41	0.39	0.12	0.19	57
			42	0.36	0.07	0.12	58
			43	0.68	0.48	0.56	540
			44	0.77	0.29	0.42	59
			45	0.61	0.32	0.42	358
			46	0.80	0.58	0.67	1360
			47	0.74	0.54	0.63	1128
			48	0.81	0.58	0.68	558
			49	0.70	0.36	0.48	136
			50	0.79	0.40	0.54	84
			51	0.58	0.20	0.29	187
			52	0.47	0.09	0.16	75
			53	0.57	0.09	0.15	47
			54	0.75	0.64	0.69	56
			55	0.95	0.55	0.69	64
			56	0.73	0.44	0.55	141
			57	0.82	0.54	0.65	239
			58	0.92	0.66	0.77	175
			59	0.73	0.43	0.54	51

			57	0.82	0.54	0.65	239
			58	0.92	0.66	0.77	175
			59	0.73	0.43	0.54	51
			60	0.52	0.29	0.37	354
			61	0.00	0.00	0.00	56
			62	0.00	0.00	0.00	46
			63	0.81	0.47	0.60	55
			64	0.67	0.29	0.40	126
			65	0.47	0.11	0.17	168
			66	0.85	0.56	0.68	71
			67	0.89	0.62	0.73	595
			68	0.80	0.60	0.69	62
			69	0.90	0.66	0.76	868
			70	0.82	0.49	0.61	47
			71	0.91	0.53	0.67	219
			72	0.86	0.57	0.68	125
			73	0.77	0.38	0.51	60
			74	0.75	0.38	0.50	287
			75	0.79	0.54	0.64	289
			76	0.38	0.06	0.10	87
			77	0.98	0.68	0.80	139
			78	0.78	0.23	0.35	61
			78	0.78	0.23	0.35	61
			79	0.58	0.28	0.38	67
			80	0.79	0.60	0.69	103
			81	0.61	0.35	0.44	295
			82	0.67	0.46	0.54	194
			83	0.41	0.15	0.22	148
			84	0.80	0.67	0.73	42
			85	0.95	0.64	0.76	108
			86	0.86	0.36	0.51	50
			87	0.64	0.29	0.40	79
			88	0.07	0.02	0.03	56
			89	0.90	0.57	0.70	67
			90	0.79	0.47	0.59	128
			91	0.95	0.62	0.75	61
			92	0.49	0.31	0.38	108
			93	0.64	0.36	0.46	75
			94	0.92	0.67	0.77	51
			95	0.45	0.13	0.21	150
			96	0.71	0.23	0.35	64
			97	0.91	0.69	0.78	163
			98	0.76	0.38	0.51	182
			99	0.69	0.45	0.54	110
			micro avg	0.77	0.47	0.59	19857
			macro avg	0.69	0.39	0.48	19857
			weighted avg	0.74	0.47	0.57	19857
			samples avg	0.58	0.52	0.52	19857
			Process finished with exit code 0				

2.	Tune the Model	Hyperparameter Tuning Validation Method	 <p>The screenshot shows a Jupyter Notebook titled 'predicting-tags' with a 'Last Checkpoint' timestamp. The code cell contains two lines: <code>print('Precision', prec)</code> and <code>print('Recall', recall_score(y_test, y_pred, average='weighted'))</code>. The output displays performance metrics for seven different classifiers: SGDClassifier, LogisticRegression, MultinomialNB, LinearSVC, and RandomForestClassifier. Each classifier's output includes Hamming Loss, Precision, and Recall values.</p> <table border="1"><thead><tr><th>Classifier</th><th>Hamming Loss</th><th>Precision</th><th>Recall</th></tr></thead><tbody><tr><td>SGDClassifier</td><td>0.40955594817646855</td><td>0.833768232725876</td><td>0.45127662788948025</td></tr><tr><td>LogisticRegression</td><td>0.8007126886965316</td><td>0.8844362489333895</td><td>0.4726292994813033</td></tr><tr><td>MultinomialNB</td><td>0.8117580939757478</td><td>0.717938689846389</td><td>0.4353628443378896</td></tr><tr><td>LinearSVC</td><td>0.80955274647807584</td><td>0.773397768682136</td><td>0.5245881762682689</td></tr><tr><td>RandomForestClassifier</td><td>0.8103153339966681</td><td>0.828788988724895</td><td>0.4885136727682357</td></tr></tbody></table>	Classifier	Hamming Loss	Precision	Recall	SGDClassifier	0.40955594817646855	0.833768232725876	0.45127662788948025	LogisticRegression	0.8007126886965316	0.8844362489333895	0.4726292994813033	MultinomialNB	0.8117580939757478	0.717938689846389	0.4353628443378896	LinearSVC	0.80955274647807584	0.773397768682136	0.5245881762682689	RandomForestClassifier	0.8103153339966681	0.828788988724895	0.4885136727682357
Classifier	Hamming Loss	Precision	Recall																								
SGDClassifier	0.40955594817646855	0.833768232725876	0.45127662788948025																								
LogisticRegression	0.8007126886965316	0.8844362489333895	0.4726292994813033																								
MultinomialNB	0.8117580939757478	0.717938689846389	0.4353628443378896																								
LinearSVC	0.80955274647807584	0.773397768682136	0.5245881762682689																								
RandomForestClassifier	0.8103153339966681	0.828788988724895	0.4885136727682357																								