




























Competition:	Sinkhole or Not																					
Kaggle Username:	Shadowless																					
Best Performing Method:	Random Forest Classifier																					
Optimized Parameter and Values																						
Parameters	Value																					
min_samples_leaf	2																					
min_samples_split	3																					
n_estimators	400																					
Max_depth	30																					
bootstrap	False																					
Criterion	entropy																					
Summary																						
An hyperparameter tuning was applied on the default decision tree classifier and the optimized parameter values was tested on the data. The result was not quite a significant improvement on the baseline; therefore, a Random forest classifier was applied using an n-estimator = 100, max_depth = 9. The result was equal to that of the optimized decision tree and an hyperparameter tuning was subsequently performed. The resulting parameter values were used on the data, and I was able to get a significant improvement on the baseline.																						
Kaggle Leaderboard																						
Public Leaderboard																						
<table><tr><td>2</td><td>Fahmida Rahman</td><td></td><td>0.89810</td><td>12</td><td>1d</td></tr><tr><td>3</td><td>ahwachal</td><td></td><td>0.89374</td><td>5</td><td>18h</td></tr><tr><td>4</td><td>Shadowless</td><td></td><td>0.89301</td><td>12</td><td>16h</td></tr></table>		2	Fahmida Rahman		0.89810	12	1d	3	ahwachal		0.89374	5	18h	4	Shadowless		0.89301	12	16h			
2	Fahmida Rahman		0.89810	12	1d																	
3	ahwachal		0.89374	5	18h																	
4	Shadowless		0.89301	12	16h																	
Private Leaderboard																						
<table><tr><td>8</td><td> 3</td><td>Connor Blandford</td><td></td><td>0.88958</td><td>11</td><td>4d</td></tr><tr><td>9</td><td> 2</td><td>Luke Ritchie</td><td></td><td>0.88927</td><td>5</td><td>17h</td></tr><tr><td>10</td><td> 6</td><td>Shadowless</td><td></td><td>0.88864</td><td>12</td><td>16h</td></tr></table>		8	 3	Connor Blandford		0.88958	11	4d	9	 2	Luke Ritchie		0.88927	5	17h	10	 6	Shadowless		0.88864	12	16h
8	 3	Connor Blandford		0.88958	11	4d																
9	 2	Luke Ritchie		0.88927	5	17h																
10	 6	Shadowless		0.88864	12	16h																
My best kaggle public score was 0.89301 and 0.8864 on the private leaderboard																						