F1 Score Comparison on both datasets 1.0 0 F1 Score 0.5 Model string-matching 0 1.0 Score 0.5 0 0 Feature Eng. Task Type