f1 of category 2 of BERT models trained with Data generated by turbo tested with data generated by davinci for Essey Set 5 used for davinci. 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0 0.0 50 100 200 400 800 1600 3200 Batch Size