

Mean F1 performance on test data

