

Average F-score across validation sets

