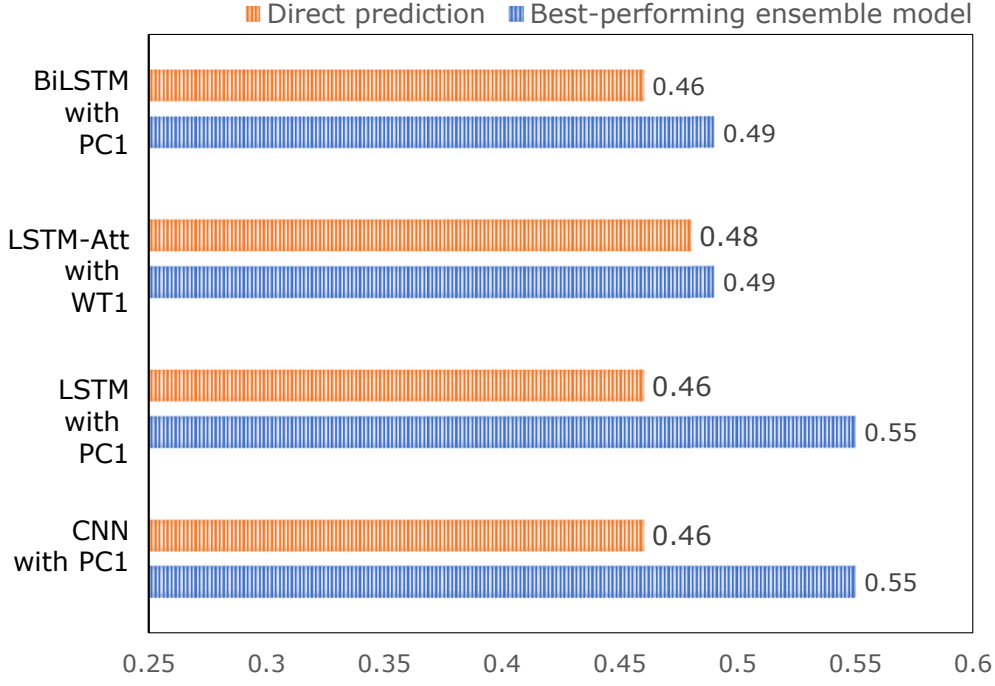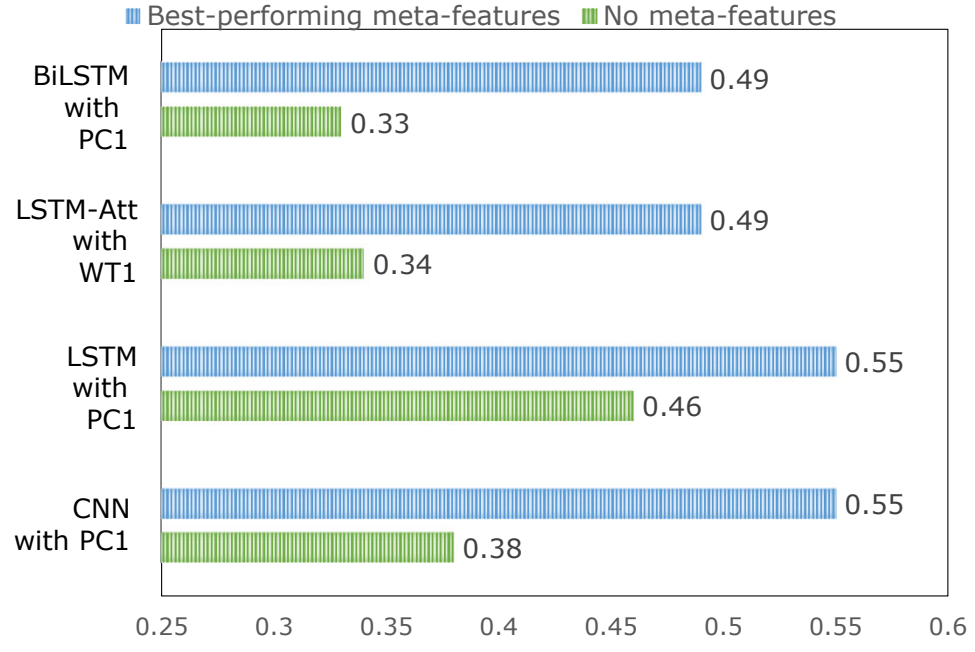IMPROVING LANGUAGE MODELS WITH ENSEMBLE MODELS: CLAFF-DIPLOMACY DATASET (F1 SCORES)