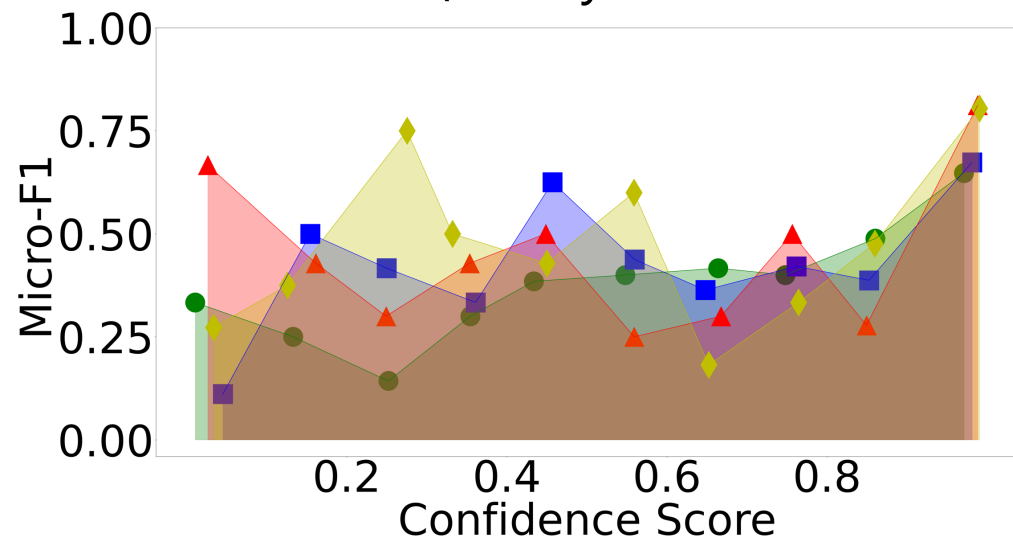
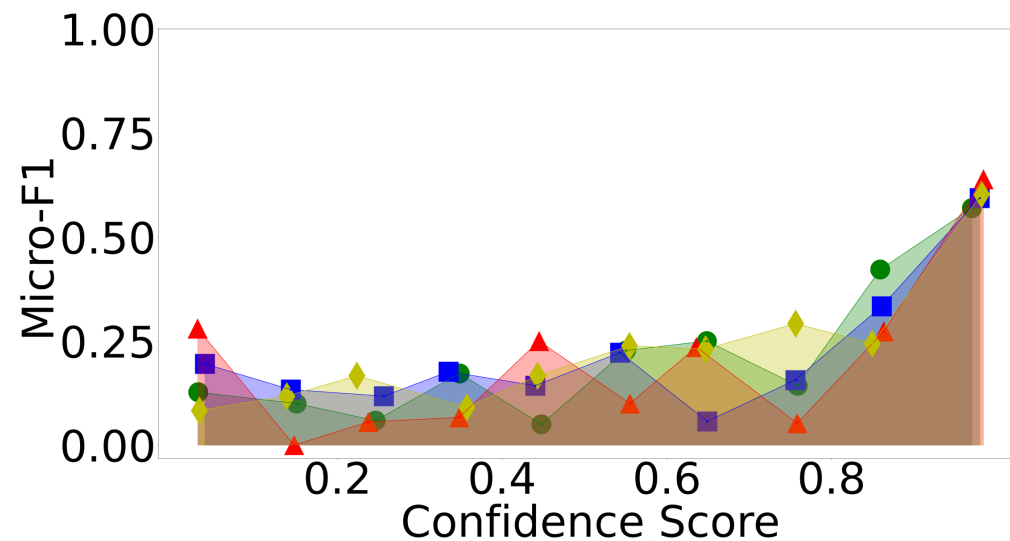


Confidence Score(Query Validation 재실험)

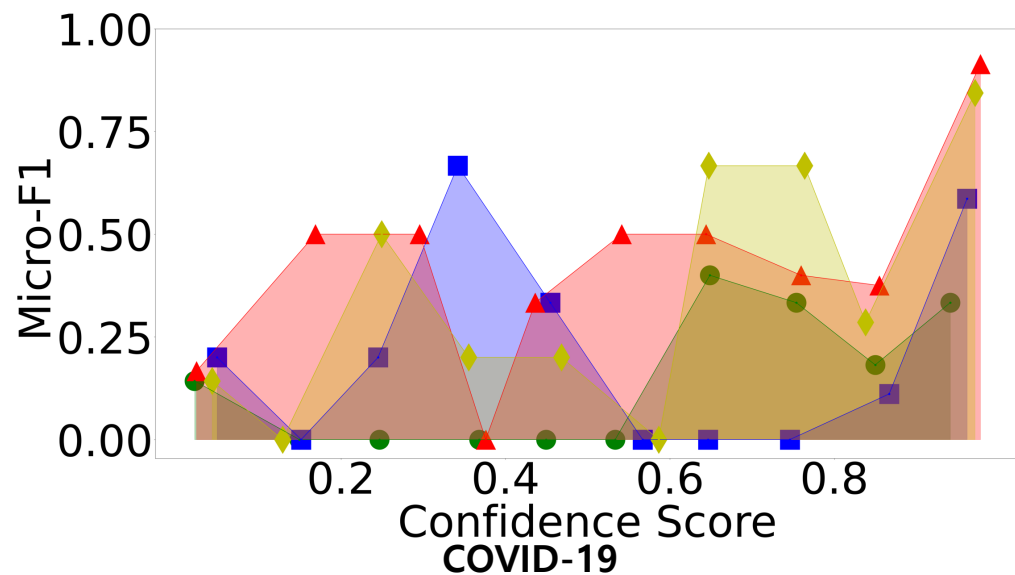


ARGOTARIO

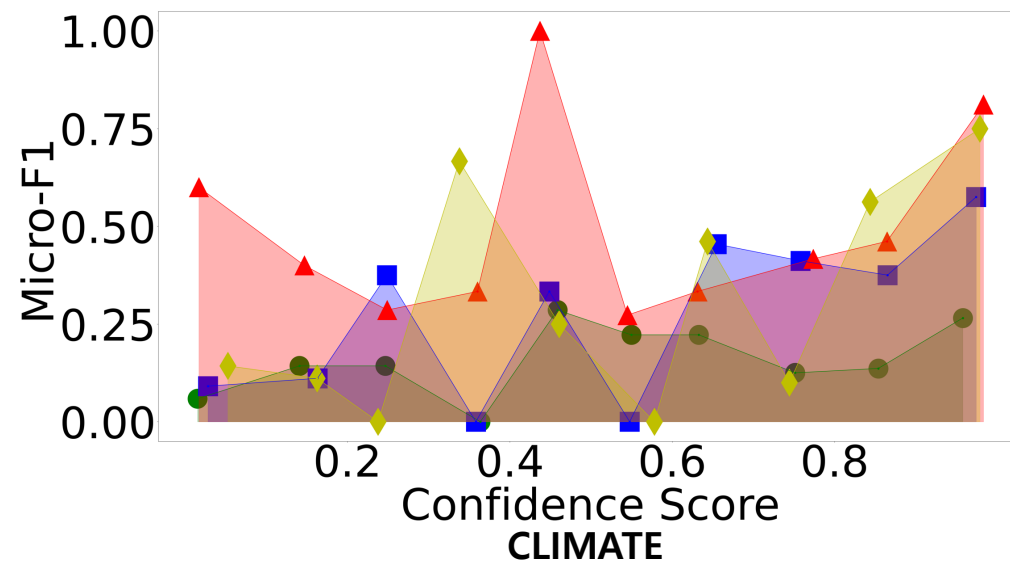
초록 : base, 파랑 : cg, 빨강 : ex, 노랑 : go



LOGIC



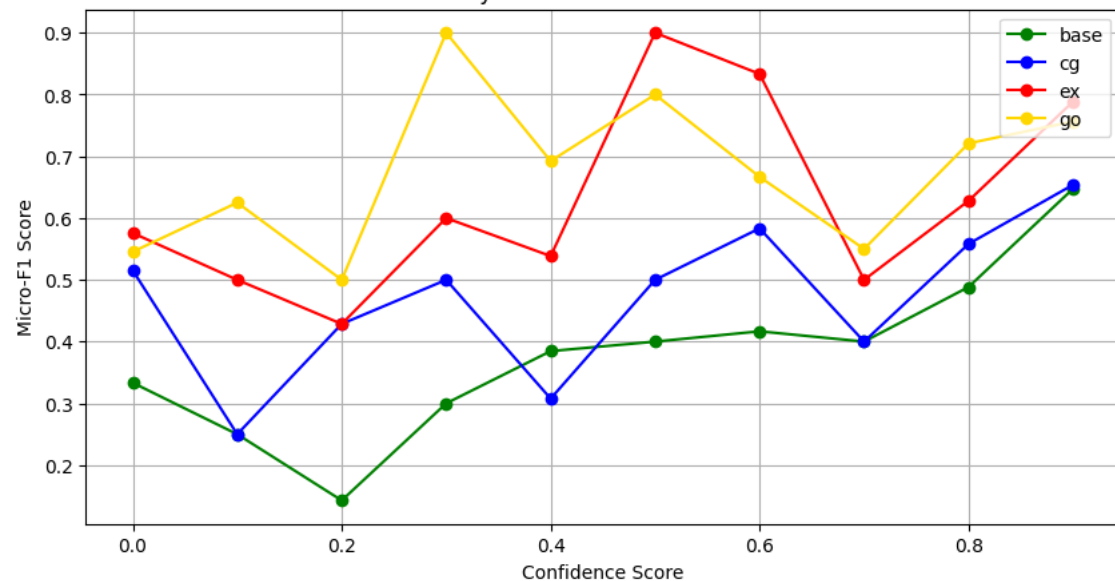
COVID-19



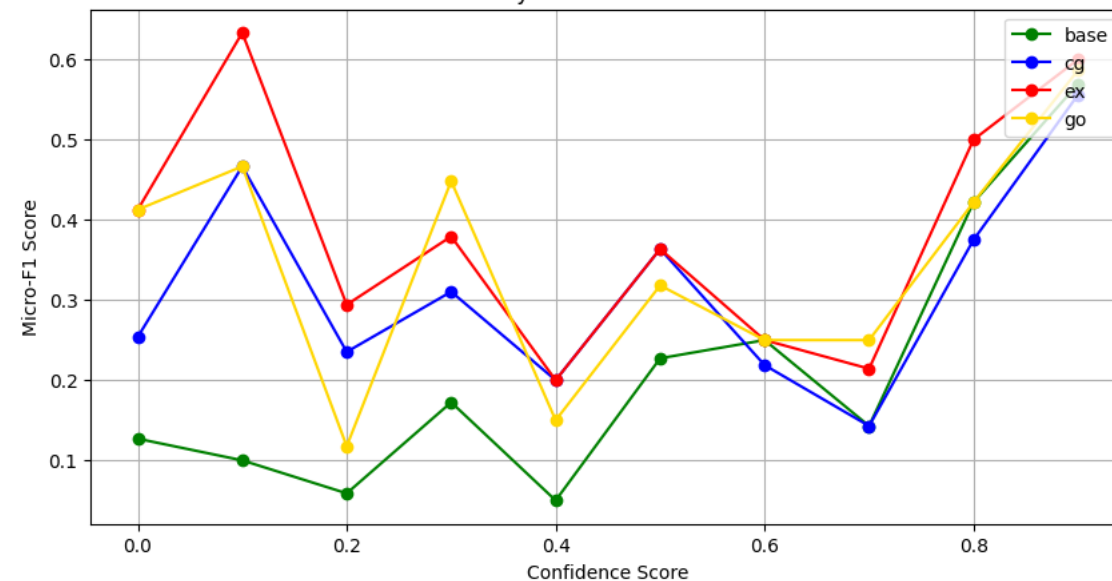
CLIMATE

Confidence score 추가실험

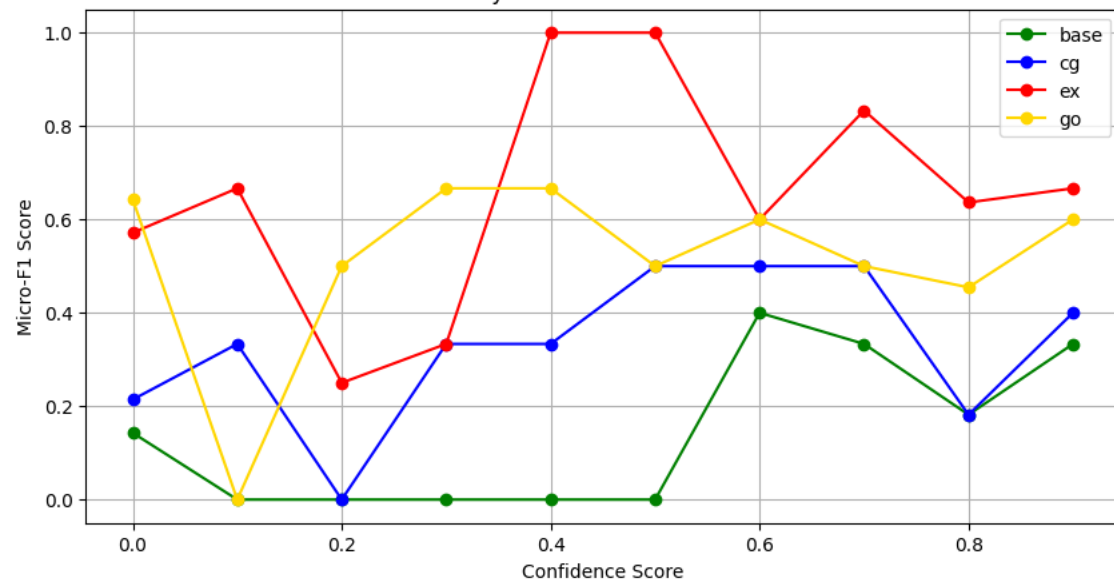
Micro-F1 Scores by Confidence Scores Across ARGOTARIO



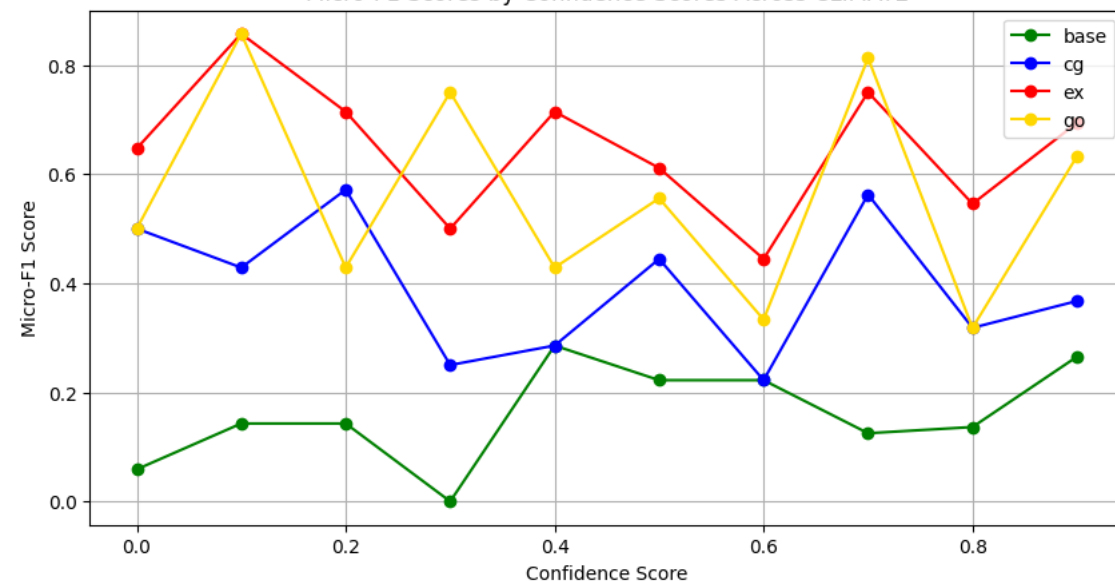
Micro-F1 Scores by Confidence Scores Across LOGIC



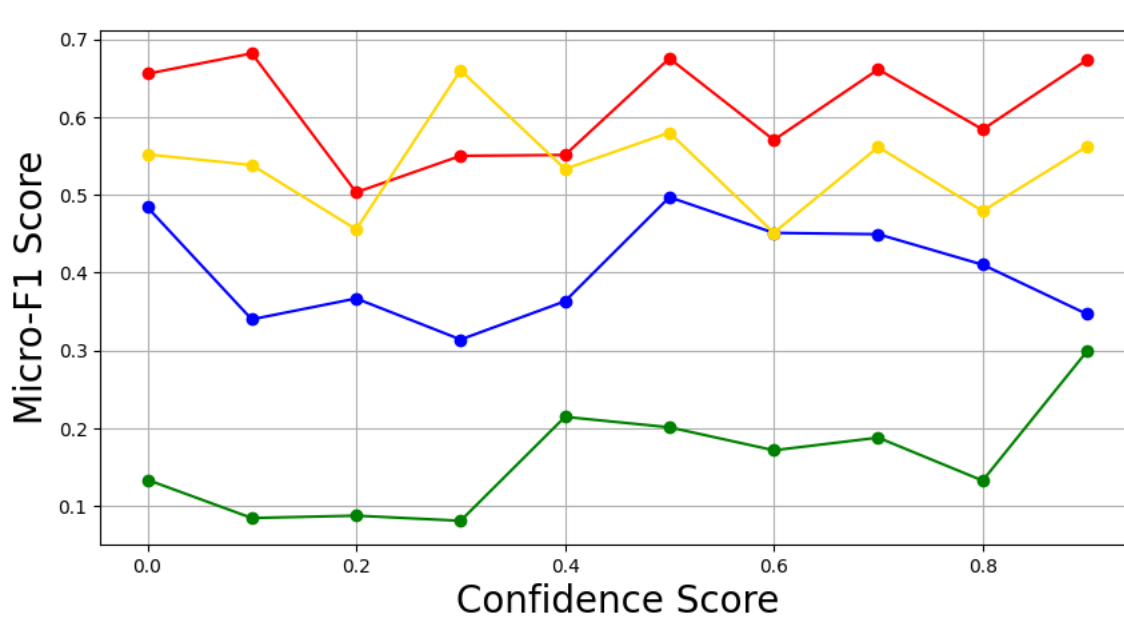
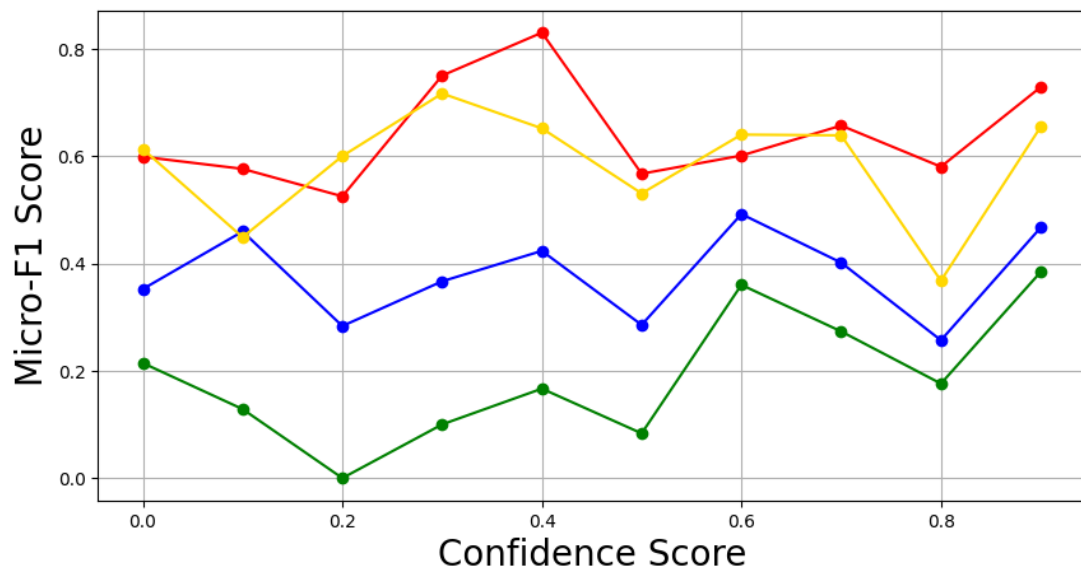
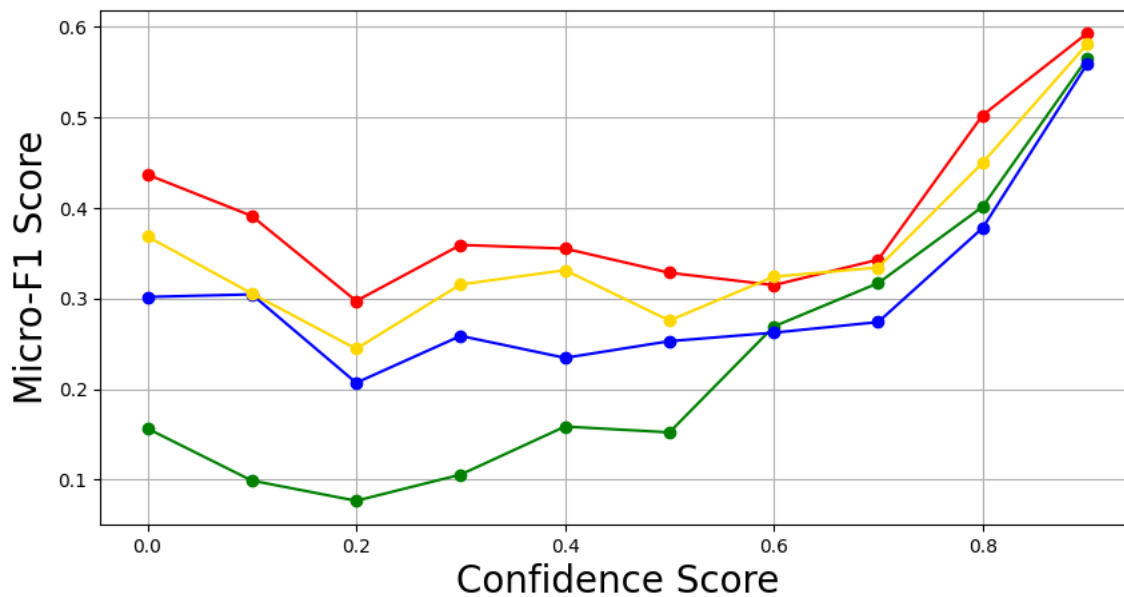
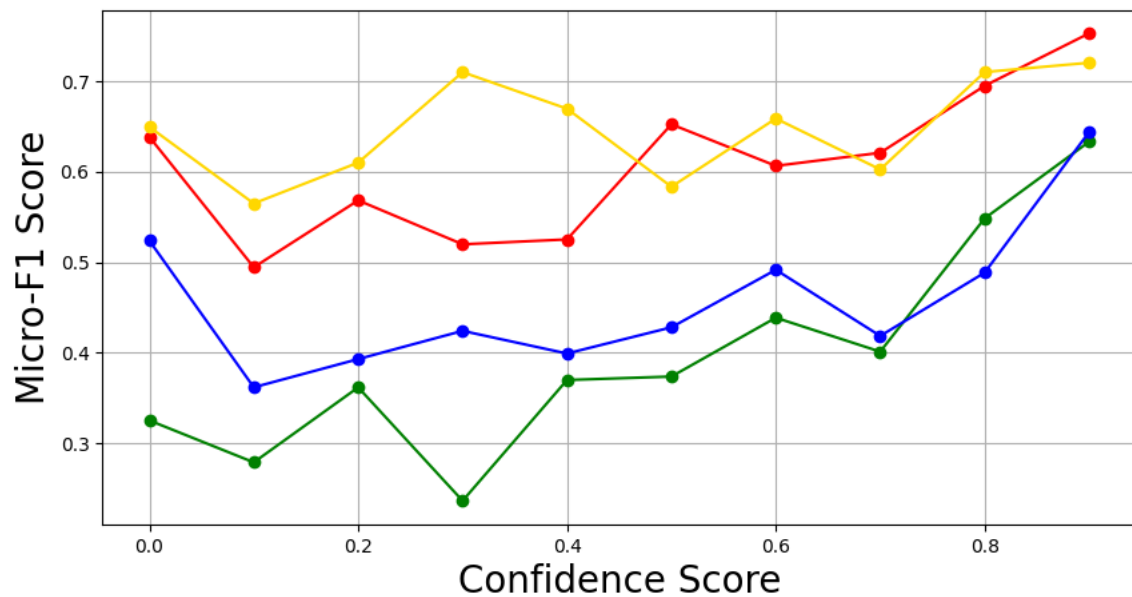
Micro-F1 Scores by Confidence Scores Across COVID-19



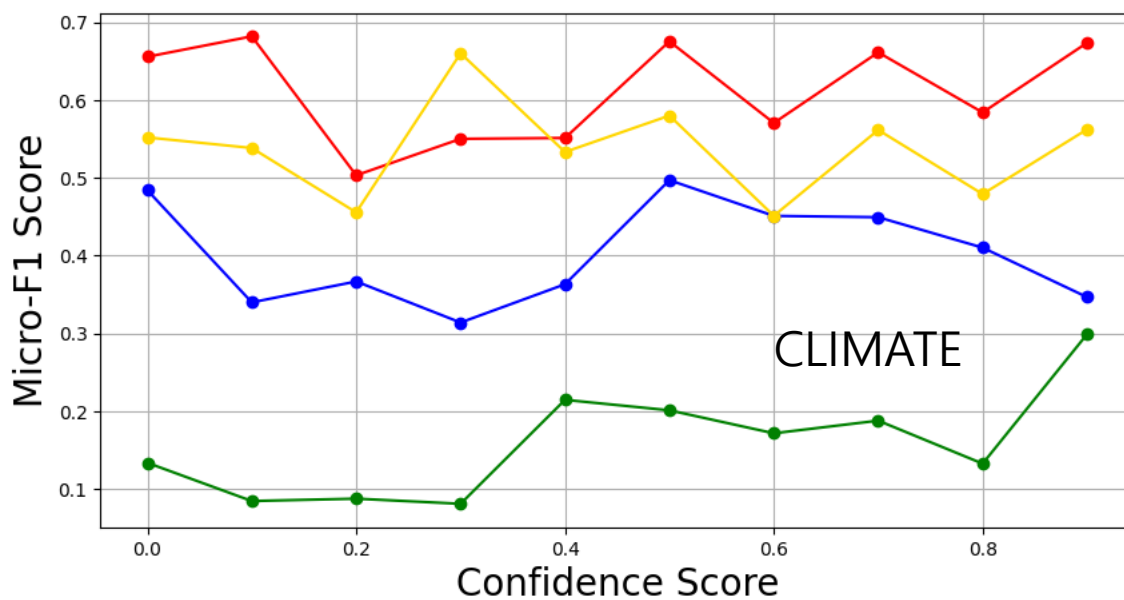
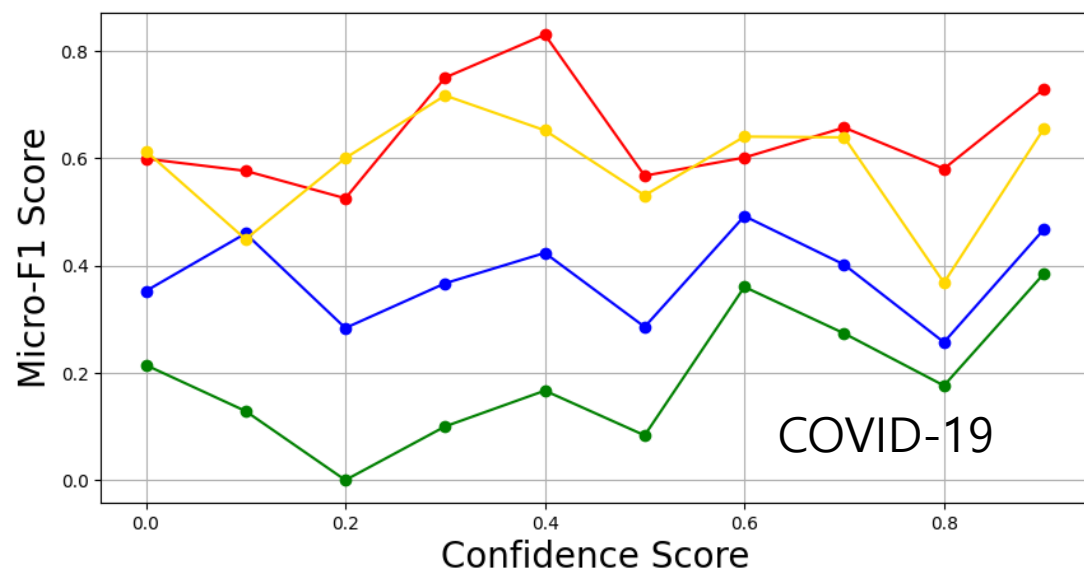
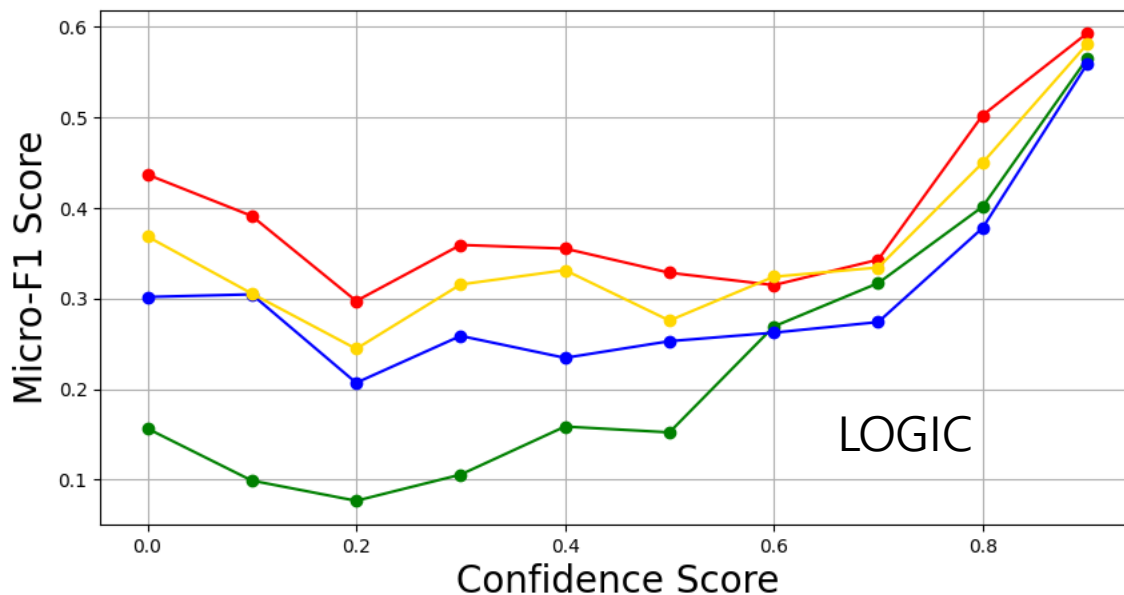
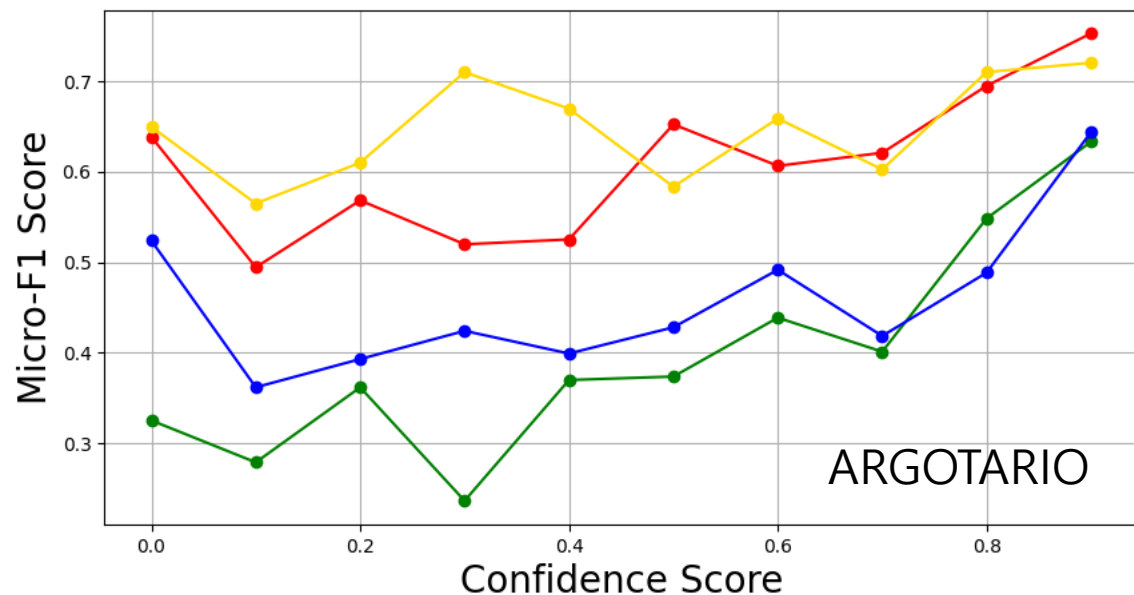
Micro-F1 Scores by Confidence Scores Across CLIMATE



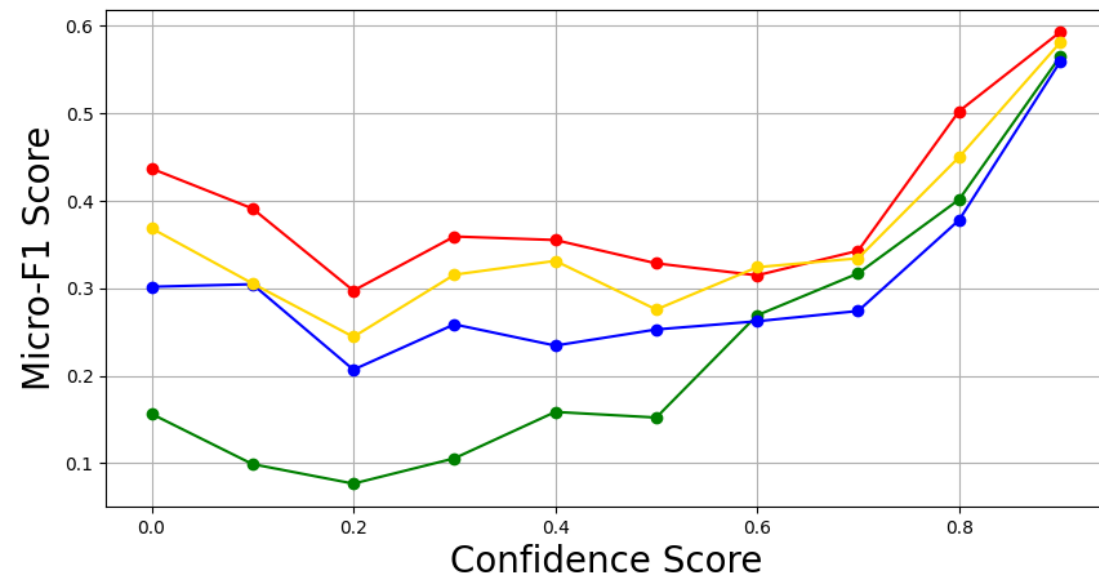
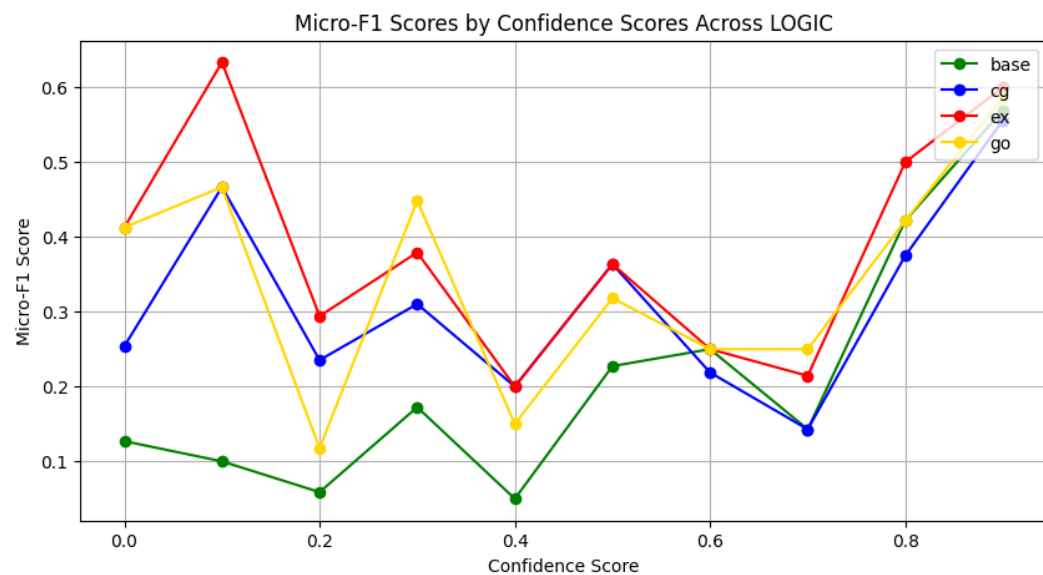
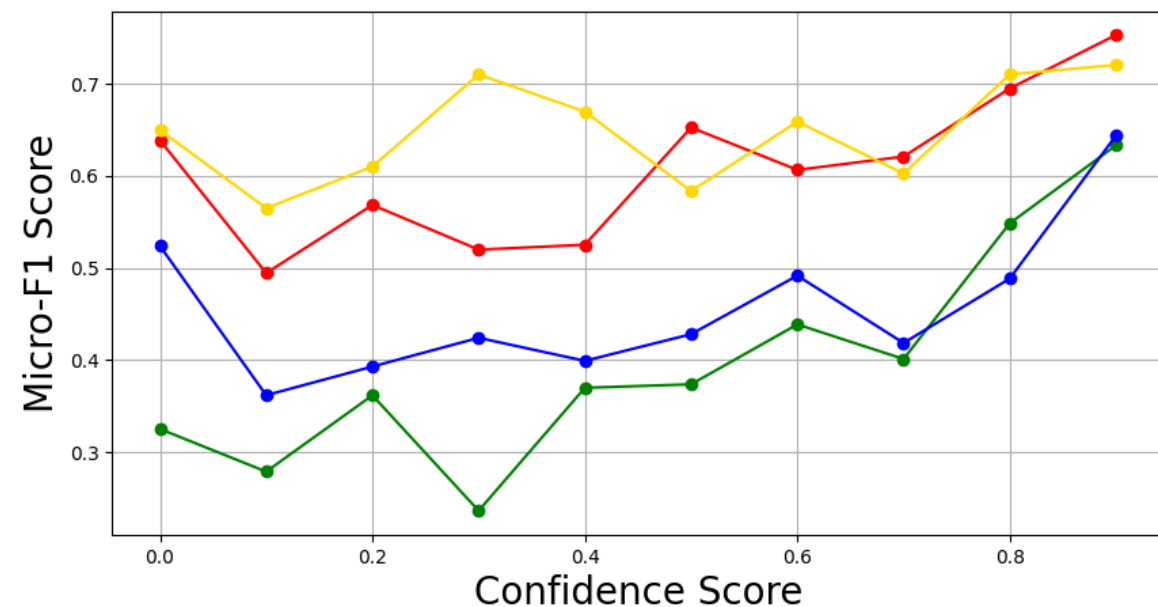
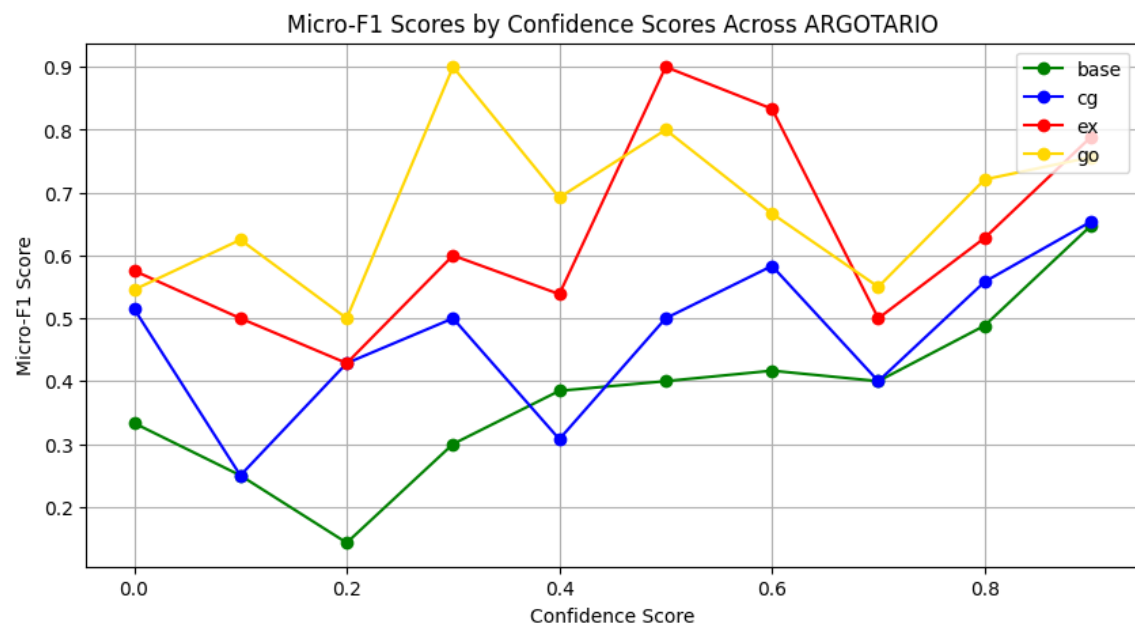
Confidence score 추가실험



Confidence score 추가실험(10회 평균)



Confidence score 추가실험(비교)



Confidence score 추가실험(비교)

