BERT for Question Answering in Biomedical Domain

Figure : Training and Testing set context availability percentage

Context availability in terms of context length in the training and testing set

Answerable examples with varying context length (Original Results)

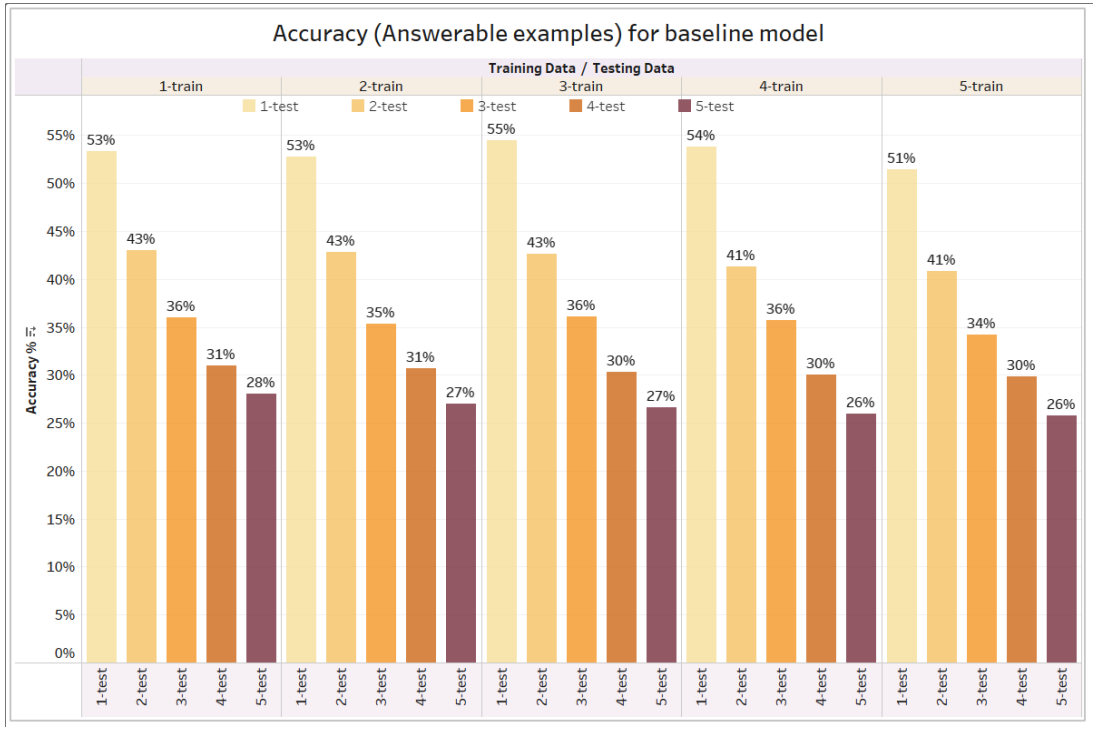

Figure : Accuracy of the baseline model with varying context length

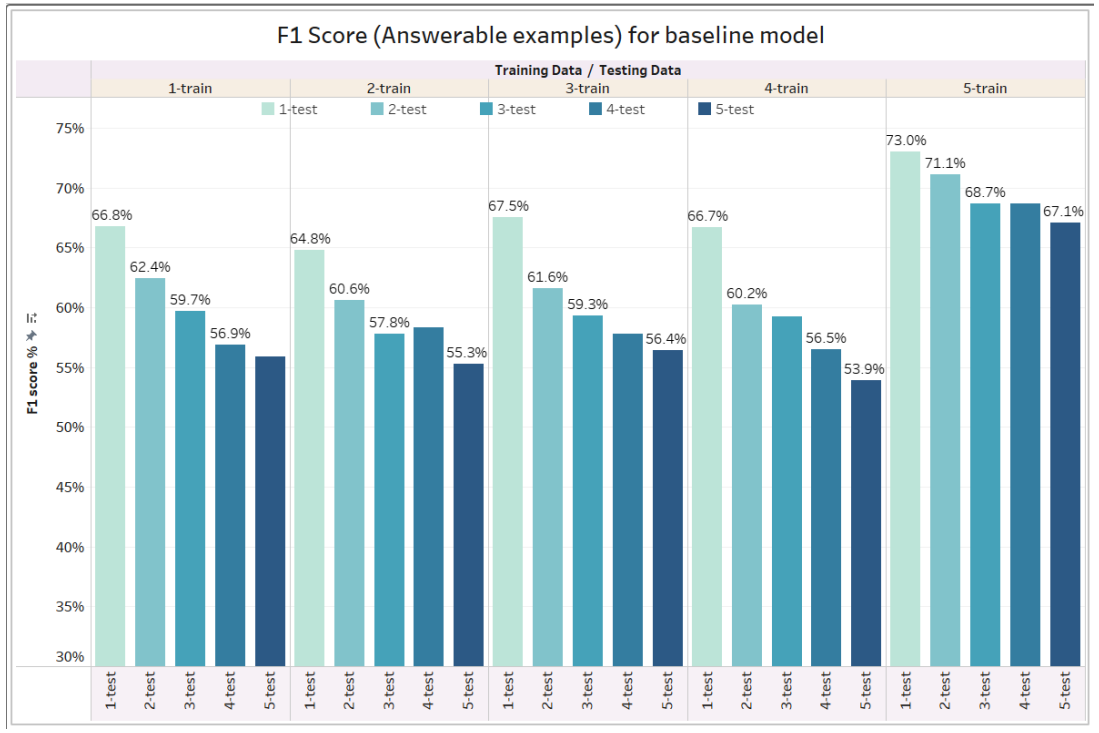

Figure : F1 score of the baseline model with varying context length

Mixed examples (Original Results)

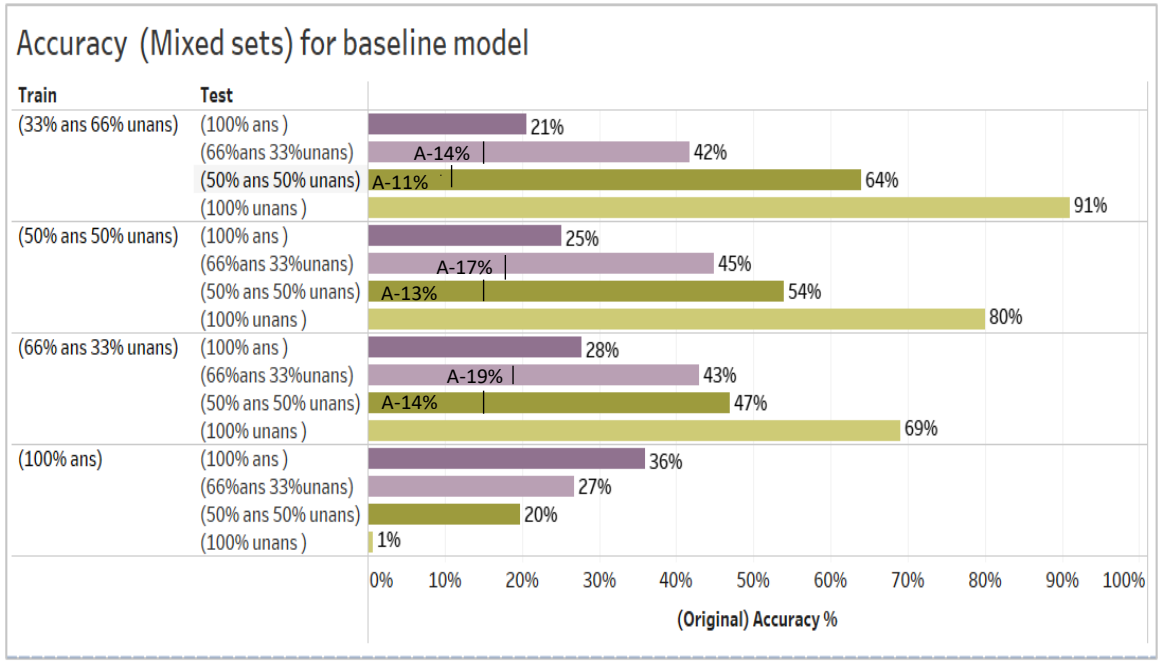

Figure : Accuracy with mixed examples

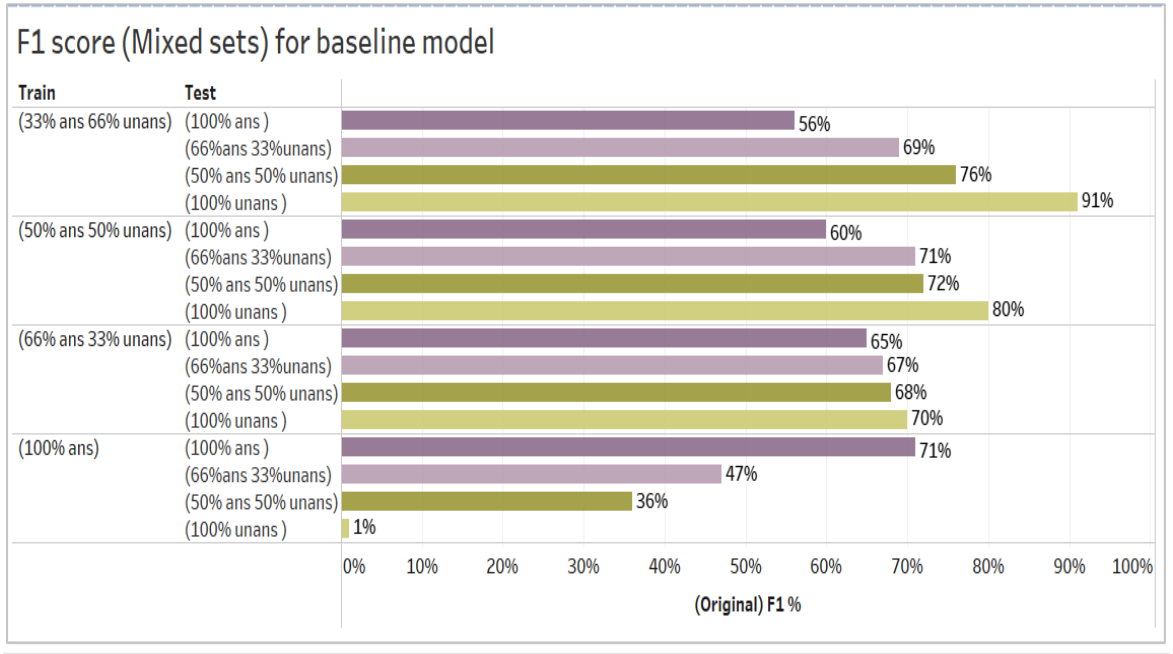

Figure : F1 score with mixed examples

Span Selection technique

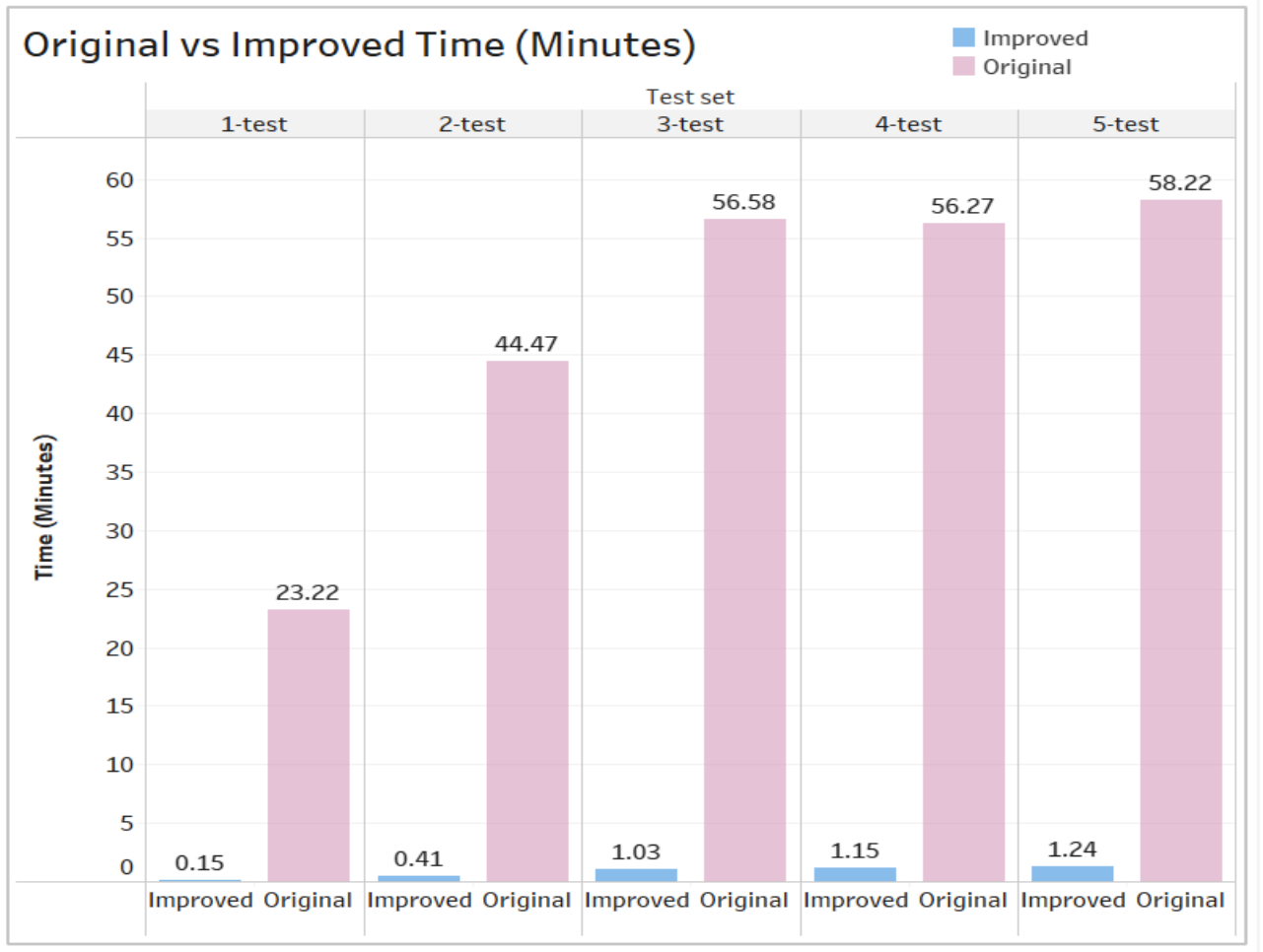

Figure : Comparison of Original and Improved Time (Examples with Varying context length)

Examples with varying context length(Original and improved)

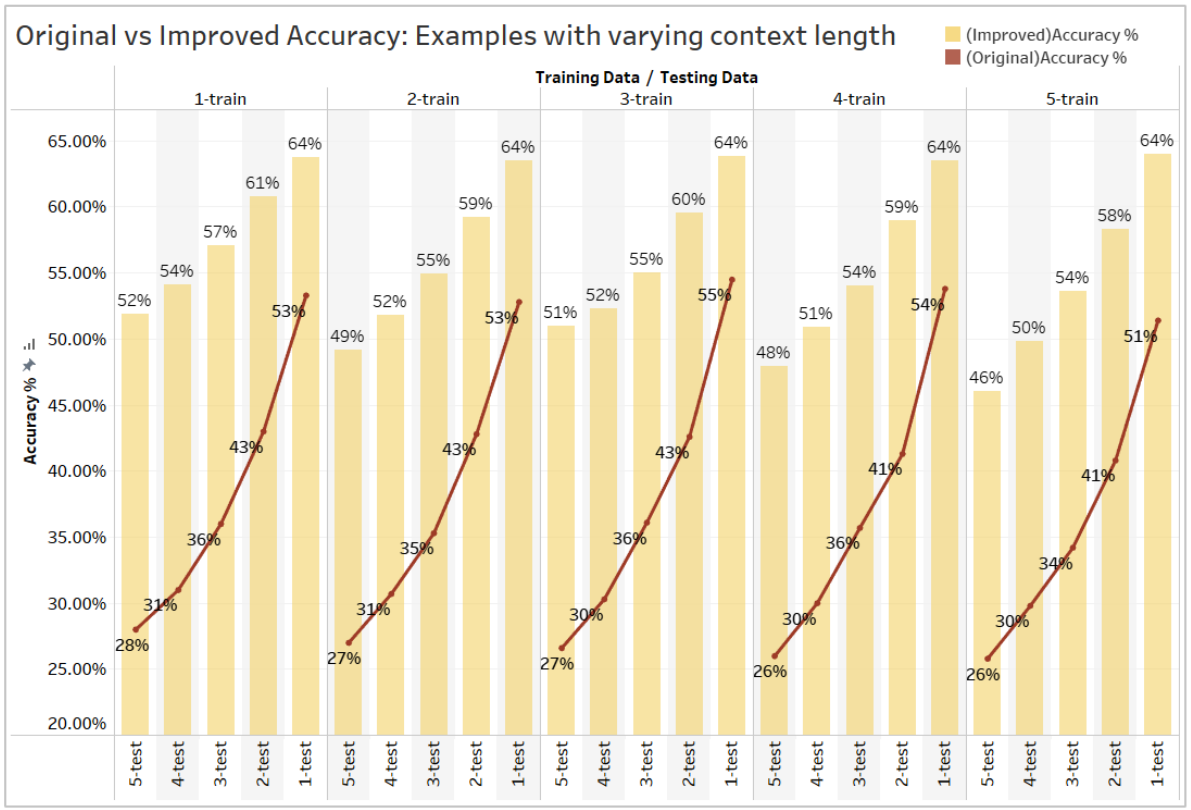

Figure : Comparison of Original and Improved Accuracy (Examples with Varying context length)

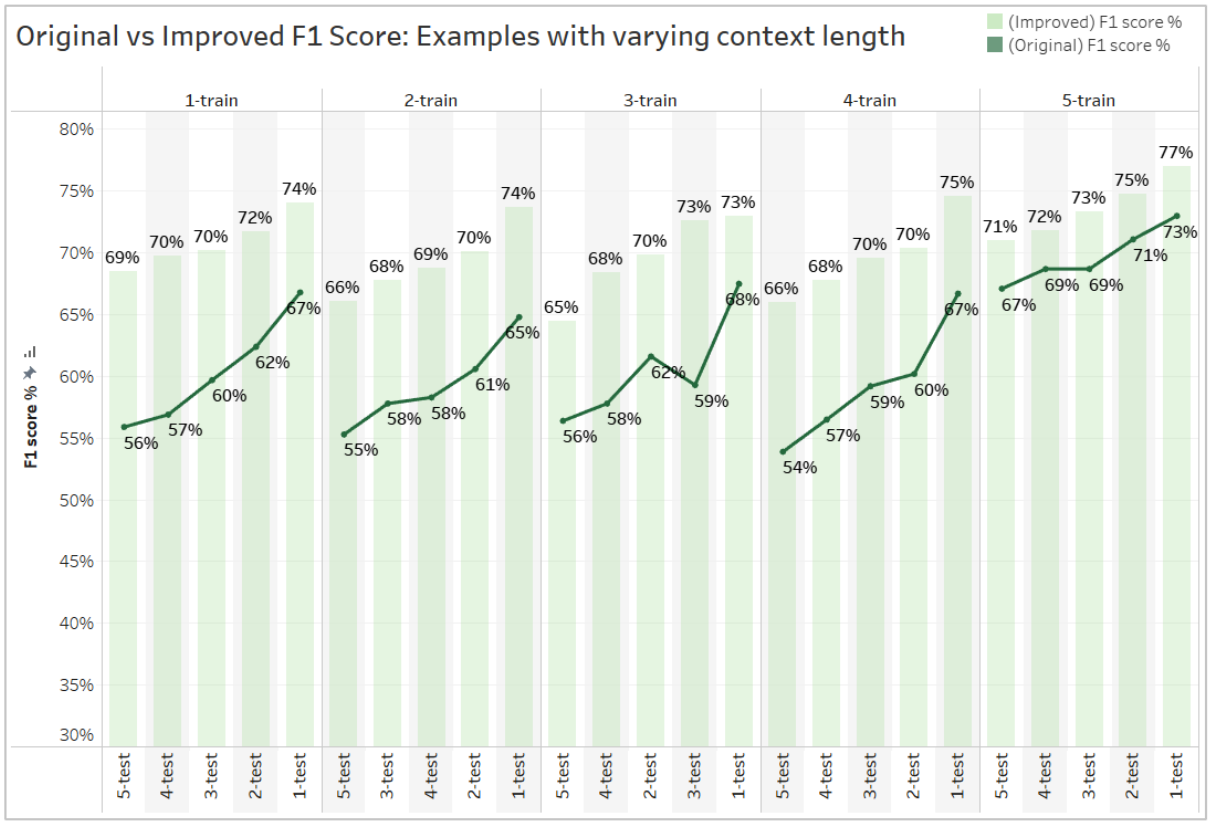

Figure : Comparison of Original and Improved F1 score (Examples with Varying context length

Mixed set (Original and improved)

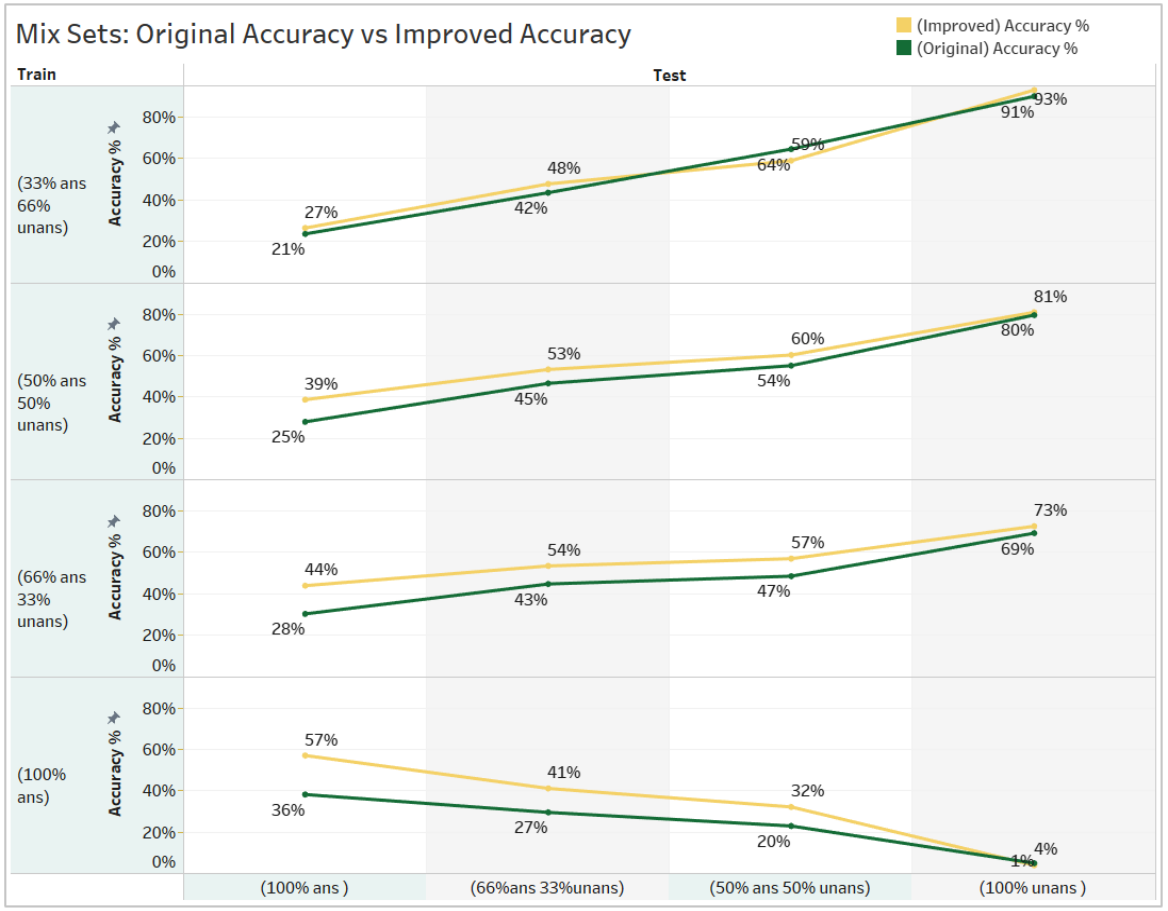

Figure : Comparison of Original and Improved Accuracy (Mixed sets)

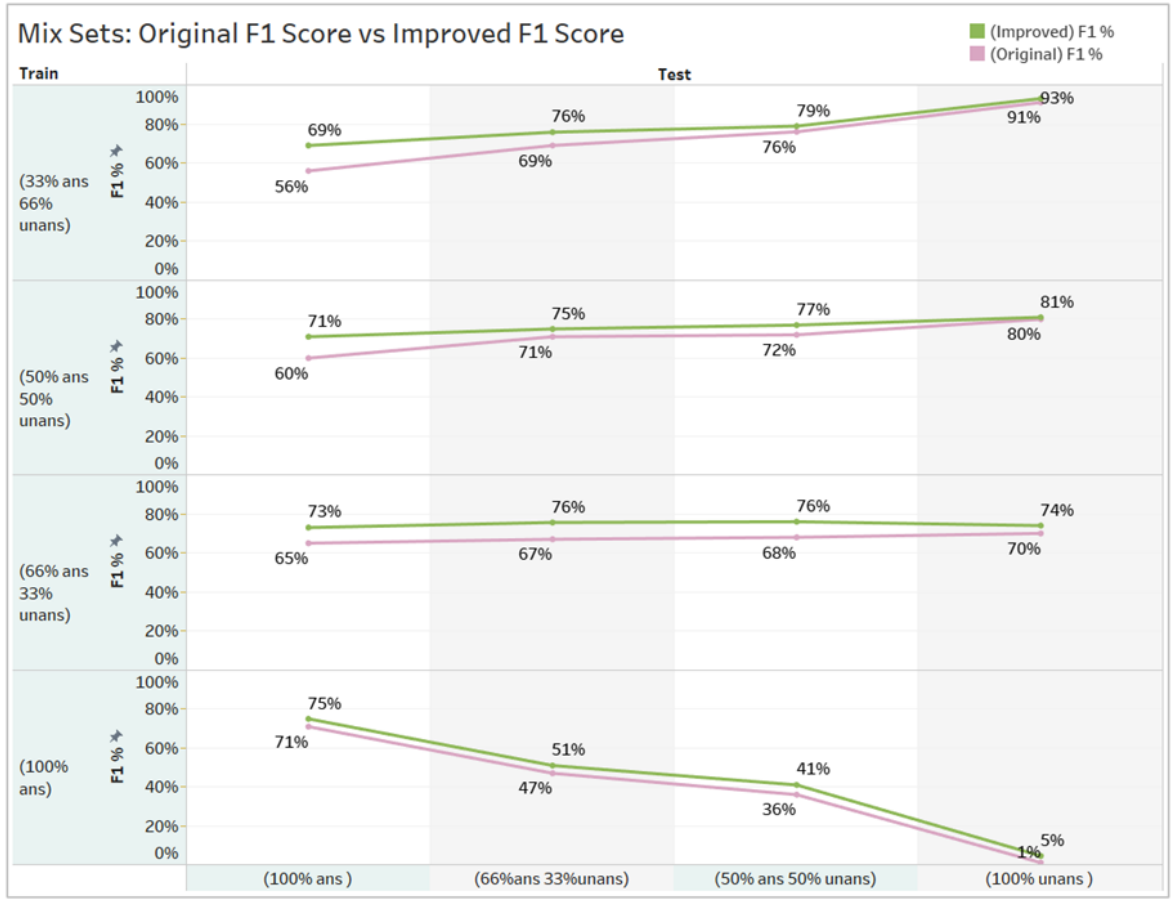

Figure : Comparison of Original and Improved F1 score (Mixed sets)