Figure 11: Comparing the language-wise performance of different models on the QA tasks.
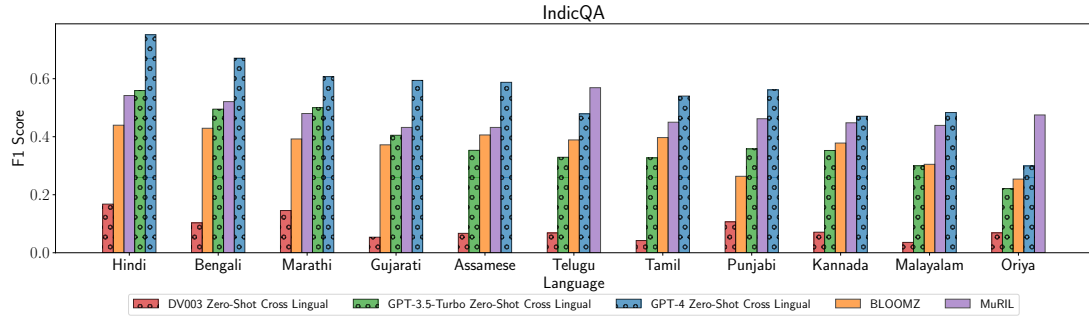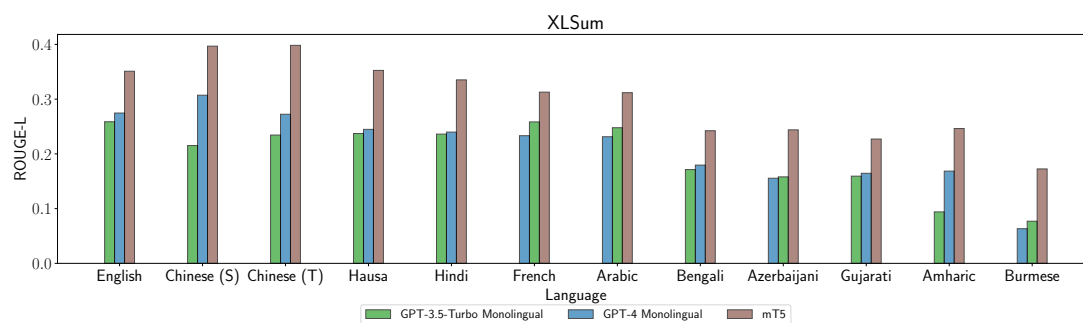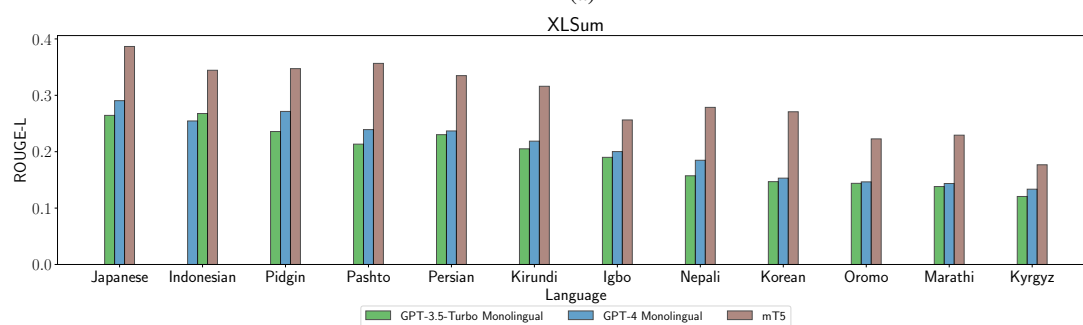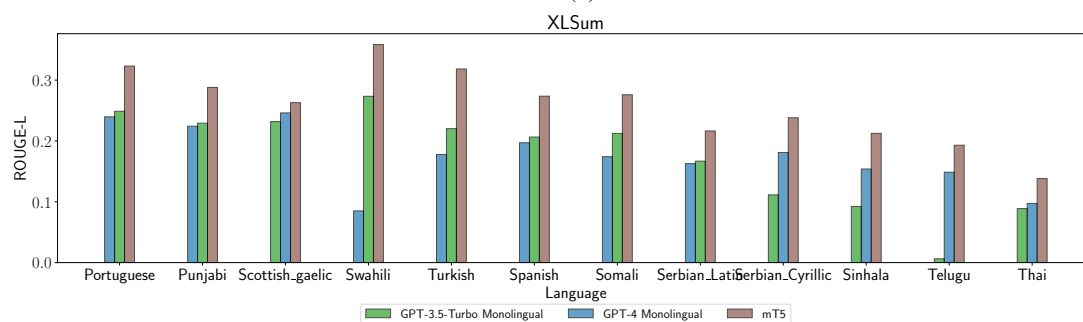
XLSum

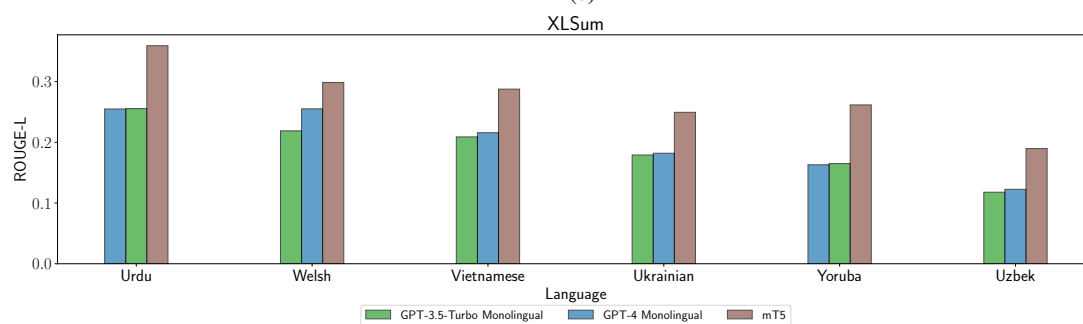(a)

XLSum

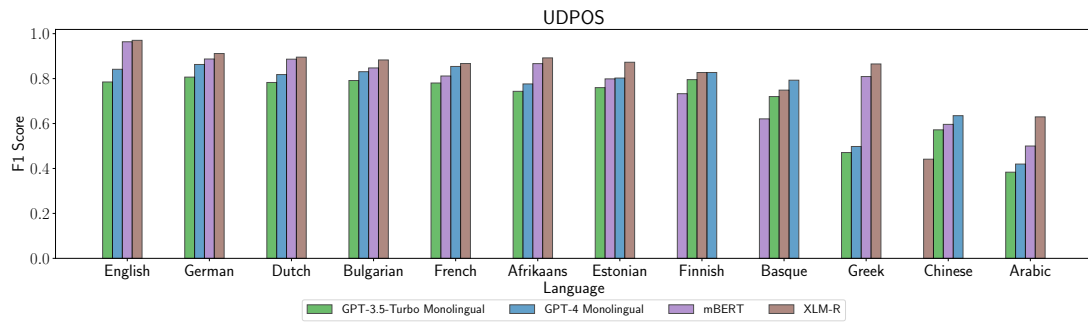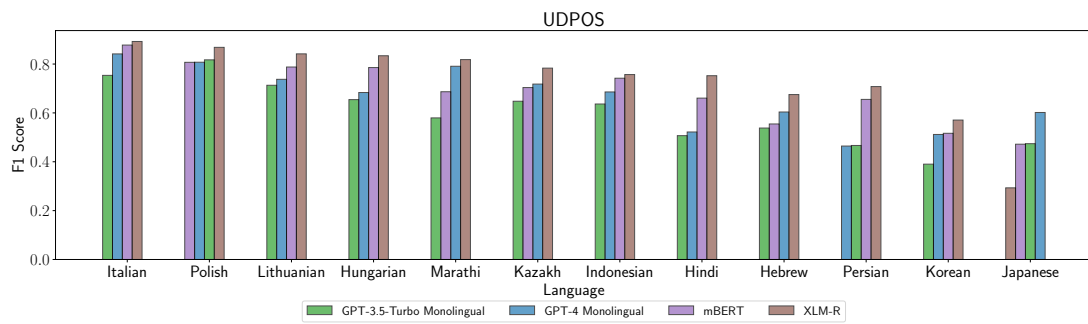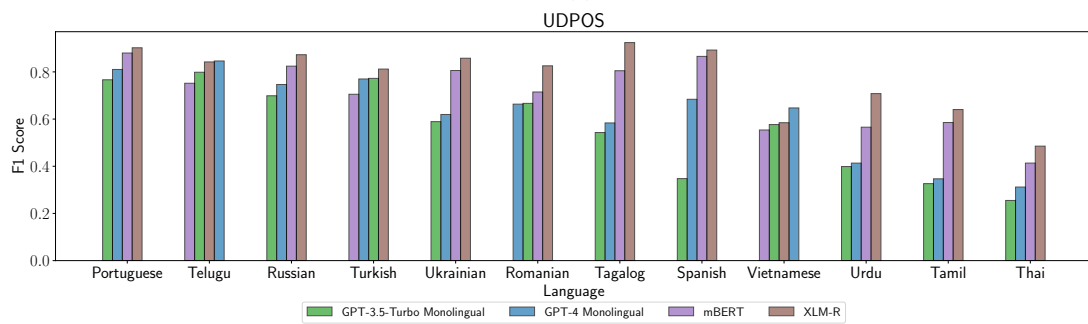(b)

XLSum

(c)

XLSum

(d)

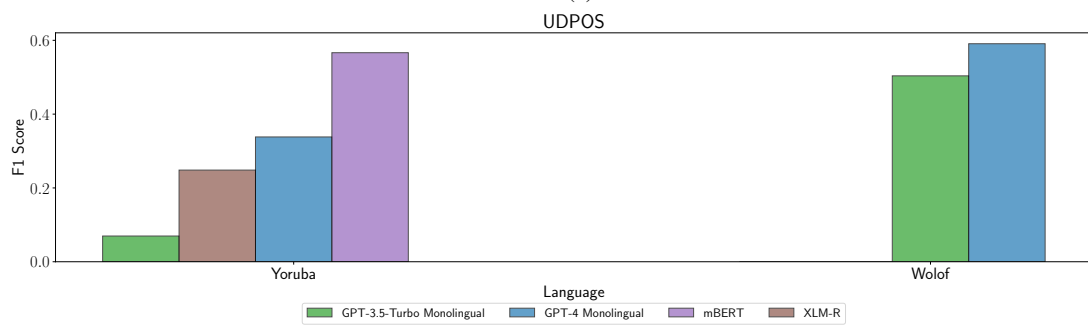Figure 12: Comparing performance of different models on XLSUM.

Figure 13: Comparing performance of different models on UDPOS