Note: Award 0 marks if model is trained during demo or code gives error in demo.

TASK 1

PERPLEXITY - 5 marks

if $10 < p < 94$, then give 5 marks
if $94 < p < 200$, then give 3 marks
if $200 < p < 300$, then give 2 marks
if $p > 300$, then give 0 marks
if $p < 10$, then give 2 marks only (penalty for using char level tokenizer)

TASK 2

ROUGE - L : 4 marks
if $r > 0.4$, then give 4 marks
if $0.33 < r < 0.4$, then give 2 marks
if $0.2 < r < 0.33$, then give 1 mark
if $r <= 0.2$, then give 0 marks

BLEU - 4 : 3 marks
if $b > 0.2$, then give 3 marks
if $0.15 < b < 0.2$, then give 1.5 marks
if $0.1 < b < 0.15$, then give 0.5 marks
if $b <= 0.1$, then give 0 marks

BERTScore : 3 marks
if $bs > 0.9$, then give 3 marks
if $0.87 < bs < 0.9$, then give 1.5 marks
if $0.85 < bs < 0.87$, then give 0.5 marks
if $bs <= 0.85$, then give 0 marks

TASK 3

ROUGE - L : 4 marks
if $r > 0.45$, then give 4 marks
if $0.4 < r < 0.45$, then give 2 marks
if $0.2 < r < 0.4$, then give 1 mark
if $r <= 0.2$, then give 0 marks

BLEU - 4 : 3 marks

if b > 0.3, then give 3 marks
if 0.16 < b < 0.3, then give 1.5 marks
if 0.1 < b < 0.16, then give 0.5 marks
if b <= 0.1, then give 0 marks

METEOR : 3 marks
if m > 0.45, then give 3 marks
if 0.26 < m < 0.45, then give 1.5 marks
if 0.2 < m < 0.26, then give 0.5 marks
if m <= 0.2, then give 0 marks