

RA1 Marking Scheme

COMP 550 Reading Assignment 3 Feedback

Read the report and use the provided marking scheme to check off points to look for. For the Coverage section, assign full marks for a good discussion, half marks for a partial discussion, zero for a missing component.

Assess linguistic quality and clarity of argumentation holistically (half marks are okay):

5/5	excellent
4/5	good: most students will score between 3 – 4 / 5
3/5	sufficient
2/5 or below	significant problems

A few potential points

A good description of the paper including the strengths and weaknesses will give you **2.5 points**

1. What are the major differences in the pre-training of mBART vs. mBERT (or XLMR) in terms of architecture, pre-training task/noise function, and pre-training data?

A good description of the differences will give you **1.5 point**, 0.5 point each for architectural, noise function and pre-training data differences.

2. Why did the authors learn a tokenizer on more languages than the languages covered in mBART25? Do you think it is a good idea? Discuss.

An answer with a good justification will give you **0.5 point** (open-ended)

3. Do high-resource language pairs benefit from mBART pre-training at the fine-tuning stage? Contrast with low-resource language pairs and discuss.

An answer with a good justification for high-resource language pairs will give you a **0.5 point**.