**COMP5623 Coursework on Image Caption Generation**

|  |  |
| --- | --- |
| Name |  |
| Student ID & username |  |

**QUESTION I [40 marks]**

**1.1 Text preparation [15 marks]**

|  |
| --- |
| Please submit your *utils.py.* |

**1.2 Extracting image features [10 marks]**

|  |
| --- |
| Please submit your *extract\_features.py* file. |

**1.3 Training DecoderRNN [15 marks]**

|  |
| --- |
| Please submit your *decoder.py* file. |

**QUESTION II [60 marks]**  
  
**2.1 Generating predictions on test data [10 marks]**

|  |
| --- |
| 2.1.1 Present three sample test images showing different objects, along with your model’s generated captions and the 5 reference captions. |

|  |  |  |
| --- | --- | --- |
| Image | Reference captions | Model generated caption |
|  |  |  |
|  |  |  |
|  |  |  |

**2.2 Caption evaluation via text similarity [30 marks]**

**(1) BLEU for evaluation**

|  |
| --- |
| 2.2.1 Report the trained model’s performance on the test set using the BLEU method, and discuss. |

|  |
| --- |
| 2.2.2 Present one sample test image with a high BLEU score and one sample with a low score, along with your model’s generated captions and the 5 reference captions. |

|  |  |  |
| --- | --- | --- |
| One sample with high BLEU score | | |
| Image | Reference captions | Model generated caption |
|  |  |  |
| One sample with low BLEU score | | |
|  |  |  |

**(2) Cosine similarity for evaluation**

|  |
| --- |
| 2.2.3 Report the trained model’s performance on the test set using the cosine similarity method, and discuss. |

|  |
| --- |
| 2.2.4 Present one sample test image with a high cosine similarity score and one sample with a low score, along with your model’s generated captions and the 5 reference captions. |

|  |  |  |
| --- | --- | --- |
| One sample with high cosine similarity score | | |
| Image | Reference captions | Model generated caption |
|  |  |  |
| One sample with low cosine similarity score | | |
|  |  |  |

**2.3 Comparing text similarity methods [15 marks]**

|  |
| --- |
| 2.3.1 Compare the model’s BLEU and cosine similarity scores on the test set and identify some weaknesses and strengths of each method. |

|  |
| --- |
| 2.3.2 Show one example where both methods give similar scores, and another example where they do not and discuss. |

|  |
| --- |
| Marks reserved for overall quality of report. [5 marks] |

*No response needed here.*