Planning (Group 30%):

Presenting a coherent plan that focuses on the practical side of the project. E.g. how we work together, our decisions, justifications and contingencies put in place before we move onto the next step. Creativity!

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Documenting progress:   * GitHub (Commit, etc.) * Gant charts (efficiency, delegation, schedule) * Monitoring |  | YOO is looking for “**creativity**” in how we document, keep a log of this project. Therefore, the criteria for ‘Planning’ is not restricted. |
| Contingencies (preventing risks):   * Regular updates to GitHub/commits |  |  |
| Justifications:   * Estimates * Resources |  |  |
| Delegation of tasks: |  |  |
| Discuss model trade-offs: accuracy vs speed |  |  |

Pre-processing (Individual 40%):

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Data visualisation:   * Descriptive statistics * Graphical visualisation |  |  |
| Very good **justification** of selection techniques:   * Explain why we used it. * Effective techniques * If technique was relevant **but not used,** then explain why. |  |  |
| Technique (not limited to):   * Transformation * Discretisation * Cleaning * Normalisation * Standardisation * Smoothing * Feature construction   Creativity! |  |  |
| Selection and application of technique is effective: |  |  |

Selecting features (Individual 40%):

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Consider techniques from following methods:   * Filter * Wrapper * Embedded |  |  |
| Dimensionality techniques (Linear and non-linear):   * PCA * Factor analysis * LDA |  |  |
| Feature construction techniques:   * Autoencoder * GANS * Literature research |  |  |
| Strong evidence of wider research and reading:   * Novel feature selection techniques and their justifications. |  |  |
| **Justification** for feature selection techniques:   * Application of FS * Effectiveness * Consideration of alternatives. |  |  |

Exploring and selecting ML algorithms (Individual 40%):

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Select candidate algorithms: 5-10 + justification |  |  |
| Establish baselines for model performance:   * Progress from simple model using initial data pipeline. |  |  |
| Justify and discuss selection strategies for algorithms:   * Effectiveness * Selection strategies of algorithms |  |  |
| Evidence of wider reading of literature:   * Novel algorithms outside the scope of lecture content. |  |  |

Refining algorithms (Individual 40%):

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Justification of selection and best configuration for hyperparameters (effectivity, insight, creativity):   * High dimensional space * Learning rate * Dropout rate * Batch size * Wider reading hyperparameters |  |  |
| Justification and consider model design components:   * Number of layers * Number of unit per layers * Loss functions * Activations * Optimisers * Drop out layer * Any novel wider components from lit. |  |  |
| Perform model-specific optimisations: reproducibility |  |  |
| Debugging model as complexity is added. |  |  |
| Discuss selection strategies for searching for best configuration:   * Trial and error * Grid search * Random search * Bayesian optimisation * Wider reading ideas. |  |  |

Evaluating model and analysing the results (Individual 40%):

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Evaluate classification performance and explain why these are chosen:   * Accuracy * Detection rate * False alarm * Type II error * MCC * TBM (Time has taken to build model) * TTM (Time taken to test model) * Wider resources to go beyond lectures. |  |  |
| Compare model’s performance with benchmarks. |  |  |
| Strong justification on evaluation methods:   * All models are fairly evaluated. * All evaluation measures are correctly calculated. * **Interpretation of results** is very good * Compared with some well-chosen sources. |  |  |
| Evidence of wider reading and wider choice of extra measures. |  | Log loss? |

Future group work (Group 15%)

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Explain where results can lead to: |  |  |
| Strengths and weaknesses: |  |  |
| Next possible steps to take: |  |  |
| Possible questions raised by results: |  |  |
| What paths were more promising? Justify. |  |  |
| What are the future plans? Collaboration.   * Unachievable in the time we spent on project * Optimistic |  |  |
| Strongly addresses question, strong evidence for future work. |  |  |

Report documentation (Group 15%):

|  |  |  |
| --- | --- | --- |
| Criteria | Complete | Comments |
| Organisation of work is:   * Well ordered * Concise * Coherent |  |  |
| **Excellent use of appendices and illustrations** |  |  |
| No mistakes (spelling, terminology) |  |  |
| Correct IEEE referencing |  |  |
| 4000 words (+- 10%) = around 850 each |  |  |
| Arial 10 point or Times New Roman 11 |  |  |
| 1.5 line spacing, 1 inch margins |  |  |
| Submitted code in .ipynb or .py format |  |  |