**Table 3.1.** Justification for the taken decisions in the methodology

|  |  |
| --- | --- |
| **Decision regarding** | **Justification** |
| Imbalanced data set | * Based on literature: Winata *et. al.* (2015) ; Charte *et. al.* (2013); Liu *et. al.* (2015); * Subsection “2.1.6. Application of the “One Vs One” problem transformation strategy for multi-label classification” of the dissertation – application of *One Vs One* approach (experiment); * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction “– application of balanced weights method(experiment); application of the stratification(experiment); * Section “2.1.10. Other classification algorithms that support multi-label classification “– application of tree-based methods with boosting(experiment); * Section “1.5. Analytical review of scientific publications on the automatic detection of learning style” (comparative studies, experiments made by other researches); |
| Correlated features in modelling | * Based on literature: [Barrera](https://www.baeldung.com/cs/author/franciscoyepes" \o "Posts by Francisco Yepes Barrera) (2021); [De Luca](https://www.baeldung.com/cs/author/gabrieledeluca" \o "Posts by Gabriele De Luca) (2020); Cook (1995); Molnar (2022) ; * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction “– application of NN; application of the stratification; * Goštautaitė, Daiva. Principal component analysis and Bloom taxonomy to personalise learning. EDULEARN19 proceedings; * Section “1.6. Data-driven approach in machine learning”; |
| Correlated features and SHAP | * Based on literature: Pedregosa *et. al.* (2011); Mase *et. al.* (2019); Aas *et. all* (2021); * Subsection “1.11.3. SHAP – model agnostic interpreatation method” – application of SHAP method; * Subsection “2.1.12. Interpretation of the predicted students’ learning style” (experiment); |
| Correlation between labels | * Based on literature: Chiang *et. al.* (2012) ; Dery *et. al.* (2021); * Subsection “Multi-label classification strategies” of the section “1.7. Supervised learning methods for classification” − application of label ranking; * Subsection “2.1.4. Application of problem transformation strategy using classfier chains for multi-label classification when labels are correlated“ − application of the Label Powerset and classifier chains for the correlated labels(experiment); * Subsection “Problem transformation methods” − KNN with label ranking (experiments (Chiang et. al.)); * Subsection “2.1.10. Other classification algorithms that support multi-label classification” − KNN with label ranking (experiments (Cause (2023))); * Table 1.9. Categorization of the multi-label classification algorythms based on the degree of correlations among labels; * Section “1.5. Analytical review of scientific publications on the topic of research conducted in the dissertation” (comparative studies, experiments made by other researches); * Chapter “Problem formulation”- classification approach for scenarios with correlated labels; * Subsection “2.1.5. Application of the “One vs all” (“One vs Rest “) problem transformation strategy for multi-label classification using different base estimators” – classification approach for scenarios with correlated labels; |
| Missing values in a data set | * Subsection „Naive Bayes or Bayes network – generative supervised machine learning model “– training with missing values; * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction” – *XGBoost* can handle missing values, but additional methods (imputation methods) have to be applied in case of NN; |
| Sparse data in a data set | * Section “1.5. Analytical review of scientific publications on the automatic detection of learning style” (comparative studies, experiments made by other researches) – XGBoost can not handle sparde data; * Subsection “2.1.6. Application of the “One vs One” problem transformation strategy for multi-label classification” – and for handling sparse data (experiment); * Subsection “2.1.9. Comparision of the classification results when applying problem adaptation approach” – comparision of the experimental results (NN); |
| Sparse labels | * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction” – poor NN performance since the labels are distributed sparsely; |
| Use of the learning style model developed in the cognitive and related theories | * Section “3.2. Developing models for student learning style identification” − clustering vs classification; * Section “2.1. Experimental Investigation (pages 147-185)” – application of supervised learning (experiments); |
| Supervised vs unsupervised | * Based on literature: Brownlee (2016); Pedregosa et al. (2011); Kim (2015); [Nakashe](file:///C:\Users\Daiva1\Desktop\Disertacija\Nakashe) (2018); * Section “1.7. Supervised learning methods for classification”; * Subsection “1.7. Supervised learning methods for classification” – supervised classification methods; * Subsection “1.10. Combining case-based reasoning methodology and Bayesian approach for students’ learning style modelling” – unsupervised machine learning modelling; * Section “2.1. Experimental Investigation (pages 147-185)” – application of supervised learning methods (experiments); * Subsection “2.1.10. Other classification algorithms that support multi-label classification” − KNN with label ranking (experiments (Cause (2023))); * Based on the experimentation conducted by [Nakashe](file:///C:\Users\Daiva1\Desktop\Disertacija\Nakashe) (2018) – K-Means application(unsupervised); |
| Generative vs Discriminative | * Based on literature: Hewitt (2018), Crossvalidated (2022); Lu *et. al.* (2019); [Rufai](https://mardiyyah.medium.com/?source=post_page-----d26def8fd64a--------------------------------) (2020); Griffiths (2019); Bishop *et. al.* (2007); [Bach](https://pubmed.ncbi.nlm.nih.gov/?term=Bach%20SH%5BAuthor%5D) *et. al.* (2017); Kim (2015); Generative models (2016); * Subsection “1.6. Data-driven approach in machine learning” – generative vs discriminative; * Subsection “2.1.9. Student learning style modeling using Gibbs sampling” – application of the generative approach; * Section “3.2. Developing models for student learning style identification” – application of generative and discriminative approaches; |
| Prediction of a single unique label combination vs predictions for each label(learning style dimension) | * Subsection “2.1.6. Application of the “One vs One” problem transformation strategy for multi-label classification” – application of the “One vs One” strategy (experiment); * Subsection “2.1.3. Application of problem transformation strategy Label Powerset methods for multi-label classification” – application of the Label Powerset method; (experiment); |
| Inherently interpretable model vs model agnotic interpretability of the model | * Based on literature: Goštautaitė et. al. (2022); Mathworks (2022), Kim (2015); [Billiau](https://sethbilliau.medium.com/?source=post_page-----b60f7d5d1fe9--------------------------------) (2021); Molnar (2022); Mase et. al. (2019); Aas et. al. (2021). * Section “1.11. Interpretation of the predictions made by learning style model” (including the subsections); * Subsection “1.11.3. SHAP – model agnostic interpreatation method” – application of SHAP method; * Subsection “2.1.12. Interpretation of the predicted students’ learning style” (experiment); |
| Epistemic and aleatory uncertainty | * Based on literature: [Indrayan](https://pubmed.ncbi.nlm.nih.gov/?term=Indrayan%20A%5BAuthor%5D) (2020); * Subsection “1.10. Combining case-based reasoning methodology and Bayesian approach for students’ learning style modelling”; |
| Conditional independence | * Based on literature: Goštautaitė (2019); Deventer (2004); * Section “Naive Bayes or Bayes network – generative supervised machine learning model”; |
| Existence of invariants/inductive bias | * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction”; * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction “ – application of NN (experiment); |
| Multi-label classification algorythms that do not require meta-estimators | * Subsection “2.1.10. Other classification algorithms that support multi-label classification” – application of Ridge classifier(experiment); |
| Size of the data | * Section “1.7. Supervised learning methods for classification” – size of the data set and the impact on the performance of the model; * Subsection “Neural networks” – impact of the size of the data set to the generalization ability; * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction“ – application of NN (experiment); * Subsection “2.1.9. Comparision of the classification results when applying problem adaptation approach” – comparision of the experimental results (NN); |
| Data normalisation, scaling | * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction” – NN performance with and without scaling (experiment); |
| Selection of hyperparameters of the model | * Subsection “2.1.5. Application of the “One vs all” (“One vs Rest “) problem transformation strategy for multi-label classification using different base estimators” – application of *GridSearch* for the selection of hyperparameters(experiments); * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction “ – application of *GridSearch* method for selection of hyperparameters for NN (experiment); |
| Handling outliers | * Subsection “2.1.1. Data collection, preprocessing and exploratory data analysis” - use of statistical methods (e.g., Z-score, IQR) or domain knowledge to identify outliers; remove outliers if this won’t lead to information loss; * Subsection „Neural networks“ - using activation functions that are less sensitive to outliers, such as the rectified linear unit (ReLU) or its variants; * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction” – NN performance with and without scaling (experiment); |
| Role of the learning rate in neural network training | * Subsection “2.1.8. Application of the problem adaptation strategy: neural network for student learning style prediction “ – application of *GridSearch* method for selection of hyperparameters for NN (experiment); |
| BR vs *One vs Rest* | * Subsection “2.1.7. Comparision of methods using problem transformation strategy”; |