Data Science and Machine Learning

Vienna University of Economics and Business
2026

Name: _____

Student ID: _____

Question	Points
1	20
2	20
3	20
4	20
5	20
Total	100

Question 1: Multiple Choice

Which statements about shrinkage methods like LASSO and Ridge Regression are true? [5 pts]

- ☐ By inducing a downward bias into the estimated parameters, they reduce the variance of the predictions.
- ☐ By introducing additional parameters for interaction terms into the OLS equation, they reduce the variance of the predictions.
- ☐ LASSO performs variable selection.
- ☐ If the sample size is low and the number of variables is high, the Ridge solution is close to the OLS solution.
- ☐ If the sample size is low and the number of variables is high, their predictions regularly outperform the OLS predictions.

Which statements about cross-validation are true? [5 pts]

- ☐ Increasing model complexity typically decreases variance but increases bias.
- ☐ Cross-validation is used to choose a tuning parameter that helps balance bias and variance.
- ☐ Cross-validation is used to estimate the out-of-sample prediction error.
- ☐ Cross-validation is used to combine the predictions of different machine learning models (ensembling).
- ☐ The prediction error will typically be lower in the test data than in the training data.

Which statements about webscraping are true? [5 pts]

- ☐ CSS selectors can help us to identify elements on a webpage we want to extract.
- ☐ Webscraping requires executing JavaScript in the browser.
- ☐ APIs are often a more stable and preferred alternative to webscraping.
- ☐ Thes existance of a Robots.txt technically prevents scraping.
- ☐ Limiting the number of requests per minute is often important to avoid overloading servers.

Which statements about map projections are true? [5 pts]

- ☐ No projection can show the earth fully undistorted.
- ☐ Conic projections suit maps with medium to large areas of the world (e.g. EU, USA)
- ☐ Azimuthal projection are especially suitable for world maps.
- ☐ Equal-area projections preserve relative sizes of regions but distort shapes.
- ☐ The Mercator projection is an equal-area projection.

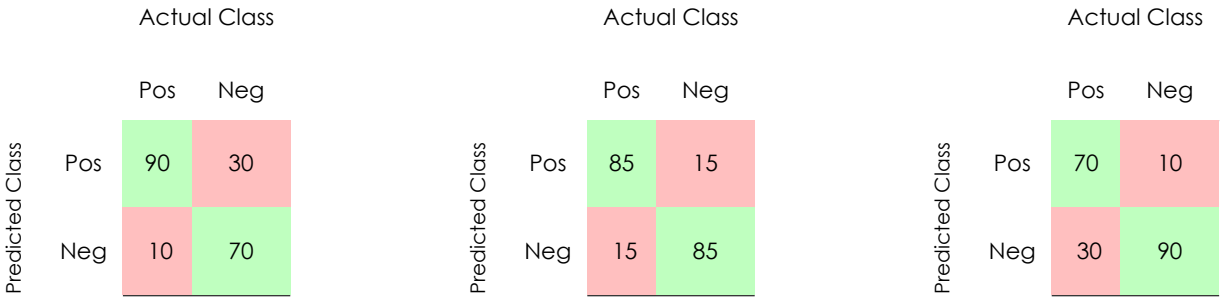
Question 2: Regression and Classification Trees

(a) Explain how a classification tree works and how it arrives at a prediction by describing the algorithm and what it does geometrically. (hint: draw a tree graph and/or a graph of the support of input variables) [10 pts]

(b) What is the curse of high dimensionality and how does it relate to trees? How can you overcome it? [10 pts]

Question 3: Classification Problems

Here are three confusion matrices obtained for three different cut-off points for the same classification problem and model.



(a) Calculate the sensitivity and the specificity for all the confusion matrices. Draw a graph of the sensitivity on the y-axis and 1-specificity on the x-axis with the three points of the respective confusion matrix and connect them to each other as well as to the extreme cases where everything is predicted into the negative and the positive category. [10 pts]

(b) What is the name of the curve that you have just drawn? Explain its importance in the context of judging the predictive performance of classification models. [10 pts]

Question 4: Cross Validation

We have learned that the predictive performance of a model can be optimized by using a process called cross validation.

(a) Explain how k-fold cross validation works. [10 pts]

(b) Explain what the bias-variance trade-off is and how it relates to cross validation. (Hint: Draw a graph of performance over complexity of a model) [10 pts]

Question 5: Data Visualization

(a) Explain what aesthetics are in context of the Grammar of Graphics. [10 pts]

(b) What are the practical advantages of plot libraries based on the Grammar of Graphics as compared to other libraries? [10 pts]