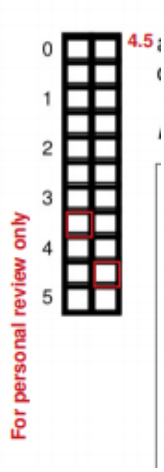
Notes on BA&ML Endterm Exam Grading and the Grade Review Process

General



- On the cover page of the review process, you will see (a) your personal details, (b) your final exam credits (points) and grade (see below) and (c) possibly some annotations for the grading. If you participated in the Analytics Cup, the grade seen in (b) is the grade for the exam only (i.e. without the AC bonus). However, the grade you see in TUMonline will include the AC bonus, so it might be different from the one listed on the cover page. We have sparingly used the annotations in (c). (Not all correctors used the annotation system, so also check any handwritten notes.) If you make a request for regrading (see below), we will use the annotations (c) to explain our decision in regrading.
- Each problem in your exam has been graded twice. The **red** comments were written by the first corrector. The second corrector has checked the work of the first corrector and may have made additional comments in **green**. Both correctors have marked your credits for that exercise by checkmarks in the margin of the page. The left column of assigned credits corresponds to the first corrector, the right column corresponds to the second corrector. The final grade for a (sub)exercise is also indicated by a red number in the margin.
- In cases of disagreement between the correctors, the credits assigned by the second corrector (**green**) are binding. Your final exam grade is thus based only on the sum of credits marked in green / in the right column. In the example on the top right, the first corrector gave 3.5 points, the second corrector overrode this decision and gave 4.5 points, which are counted toward your total points on the cover page.

- **If you used the pre-printed blackened exam pages,** (or if your scan was missing some QR codes in the corner of a page), then the grading scale that you see on the page may be misaligned, and the 'checkbox' for the assigned points may be in the wrong place. Look for the little number printed in the margin (example on the right). This indicates the final number of points you were given.
- If you believe that an exercise was not graded correctly, you can make a request for regrading of the specific (sub)problem via the TUMexam Grade Review process. In this case, a TA will review your request and grade the exercise again. **Make sure you read and follow all information in this document before making a request. Failure to do so will result in your request being denied without further explanations.** Note that, as a consequence of any request, we will completely regrade the entire (sub)problem, which may also result in the *loss of* points, not only improvements.
- To make a request: (1) carefully read the information about the grading scheme and common mistakes for the problem, which are provided below. (2) If you still believe the grading is incorrect, use the given space (500 characters) to make a *clear and concise* argument about the mistake in grading, simply writing "I believe I deserve more points." is not enough.

Grade Table

- **1.0**: ≥ 74 points
- **1.3:** ≥ 70
- **1.7**: ≥ 66
- **2.0:** ≥ 62
- **2.3:** ≥ 58
- **2.7**: ≥ 54
- **3.0:** ≥ 50
- **3.3:** ≥ 46
- **3.7:** ≥ 42
- **4.0:** ≥ 38
- Failed**:** ≤ 37.5 points

Half points are **not** rounded up to the next grade. (Example: With 65.5 points, your grade is 2.0)

Problem 1 – Regression Analysis

There were a total of **16 points** to be reached. Thereby, 1 point was given for each correctly presented argument/ calculation.

The following points summarize commonly made mistakes that led to a reduction of the points:

- Confusing dependent and independent variables led to 0 points.
- While interpreting the coefficients of a linear regression model, **not** noting that the interpretation is only valid if all other factors are held constant reduces the amount of granted points by 0.5.
- The F-Test indicates whether there exists a linear relationship between the dependent variable and at least one independent variable but is not sufficient to measure the explanatory power of the model. Thus, it is **not** a suitable answer for sub-exercises b.3 and b.5.
- In sub-exercise c.2, you should insert the provided values into the regression formula and **solve** it for the missing value.

Problem 2 – Data Preparation

There was a total of **16 points** to be reached.

- 0.5 points were given for each correct scale of measurement (Total: 4.5)
- 0.5 points were given for each correctly identified problem (Total: 4.5)
- 1 point was given for each correctly identified solution if it was appropriate. (Total 7 points.)
  - When the same solution was needed for multiple columns, the points were only given once.
  - Partial credit was given to solutions that are adequate, but are clearly not the best possible solution given the information you had access to. (Example: removing a column with many missing values rather than imputing it.)

The following were common mistakes that led to *no points* being awarded for the relevant column:

- The *statistical scale of measurement* is not the same as the *data type* of a variable. The former describes its semantics, the latter describes its representation in computer memory.

- Nominal independent variables do **not** need to be converted from *character* to *factor* data type (i.e. One-Hot-Encoded) in R, in order to use them in a Logistic Regression. This is only necessary for the **dependent** variable.
- Linear models can handle differently scaled input features (because each will be multiplied by its own coefficient).There is no need to standardize all variables to the same order of magnitude.
- Some interval/ratio (i.e. continuous) variables had highly skewed distributions. This does **not** automatically indicate the presence of outliers. It also does not warrant discretizing and binning such variables, a log-transformation is a better choice.
- Correlations clearly below 90% do not constitute multicollinearity.
- When performing *statistical inference* (as in this exercise), over-/undersampling of the data set or stratification of the target variable are not appropriate. These are techniques used in supervised learning, which was NOT present in this exercise.

Problem 3 – Naïve Bayes

Part a.) 1-Rule: A total of **5 points** were to be reached:

- 1 point was given for explaining how to properly handle the "Timestamp" column.
- 3 points were given for selecting the correct attribute and deriving the correct classification.
- 1 point was given for discussing a limitation of the 1-rule.

Common Mistakes:

- Converting "Timestamp" into a numeric attribute does not resolve the problem that it represents a unique identifier for each instance.
- 1-rule != Naïve Bayes. You need not compute conditional probabilities.
- The error rate for an attribute is neither the sum nor the average of the attribute error rates for "Spam=Yes" and "Spam=No".

Part b.) Naïve Bayes: A total of **10 points** were to be reached:

- 2 point were given for a correct discretization of "#Words".
- 5 points were given for deriving correct conditional probabilities.
- 2 points were given for providing final probabilities for classification.
- 1 point was given for stating the correct classification.

Common Mistakes:

- "#Words" should be discretized in two *equal-sized* bins.
- Zero-frequency correction is *not* necessary here. Applying the zero-frequency correction unnecessarily distorts the conditional probabilities and results in partial credits.
- You were asked to provide final (normalized) probabilities for classification, intermediary results were thus not sufficient.
- Calculation errors lead to point deductions for the respective probability and partial deductions for the final probabilities..

Part c.) Classification Problems: A total of **3 points** were to be reached:

- 1.5 points for recognizing the zero-frequency problem and suggesting a Laplacian correction.
- 1.5 points for recognizing a violation of the independence assumption and suggesting to disregard the additional attribute.

Common Mistakes:

- Violation of independence and violation of equal importance refer to the same root cause (the additional redundant attribute). Only partial credits were awarded for mentioning the second violation.
- You were asked to argue how you would handle the occurring problems. Merely stating problems only earns partial credits.

Problem 4 – Evaluation

Part a.) ROC Curve: A total of **8 points** were to be reached:

- 1 point was given for correct axes descriptions.
- 4.5 points were given for correct points on the curve.
- 1 point was given for marking the point corresponding to the cutoff value.
- 1.5 points were given for assessing the cutoff value.

Common Mistakes:

- Wrong points of the curve.
- Correct shape of the curve, but incorrect coordinates of the points. Please note that the ROC curve always terminates at (1,1) and the step sizes depend on the number of positives / negatives, respectively.
- Missing or incomplete assessment of the cutoff value.

Part b.) Colleague Statement: A total of **4.5 points** were to be reached:

- 1.5 points were given for recognizing that the colleague evaluates the model on training data, indicating overfitting.
- 1.5 points were given for recognizing that the lift curve cannot be constant at 2.
- 1.5 points were given for recognizing that the lift curve of our classifier is not monotonically decreasing.

Common Mistakes:

- Stating that the definition of a “best” classifier is not clearly defined is correct, but the colleague’s classifier does in fact perfectly classify the instances. The problem is the overfitting and evaluation on training data. Therefore, only partial credits were awarded for such statements.
- Stating that the definition of “dominating” the random classifier is not clearly defined was treated analogously.

- Stating that any classifier dominates the random classifier and that no lift curve can be constant at any value are both incorrect.

Part c.) Calculating Accuracy: A total of **5.5 points** were to be reached:

- 2 points were given for a correct idea / concept to approach this problem.
- 2.5 points were given for correct calculations.
- 1 point was given for providing a correct final accuracy.

Common Mistakes:

- Incomplete calculations, e.g. not making use of TP+TN+FP+FN=N, results in partial credits.
- From TP/(TP+FN) = 4/7 does not follow TP = 4, FN = 3!! This is incorrect.
- TN != FN, i.e. we cannot directly compute any variable directly, but we need to solve a linear equation system.

Part d.) (Dis-)Agreeing with Statement: A total of **3 points** were to be reached:

- 1.5 points were given for correctly agreeing / disagreeing.
- 1.5 points were given for a correct reasoning.

Common Mistakes:

- An example is not sufficient to prove a statement (only to disprove).
- Shuffling around formulas without clear purpose is not a sufficient reasoning.

Problem 5 – Dimensionality Reduction

Last subproblem (multiple-choice) gave 3.0 points for correct answer and 0 points otherwise.