03 Analysis - Supervised Machine Learning Readiness



Machine Learning Model Handbook

Grading Rubric

| Part 1: Problem FramingExercise 1Which scientific question should be answered? |
| --- |
| **Score** | **Criteria** |
| 5 - Excellent | * Clearly and comprehensively defines a scientific question, aligning it with the goals and context of the scenario.
* The question is appropriate for the type of machine learning analysis needed (classification or regression).
* The problem statement is well-structured, specific, and demonstrates a deep understanding of the machine learning scenario.
 |
| 4 - Proficient | * Defines a scientific question with good clarity and alignment to the goals and context of the analysis.
* The question is appropriate for the type of machine learning analysis needed (classification or regression).
* Minor improvements in specificity or structure could enhance the statement.
 |
| 3 - Satisfactory | * Provides a reasonable scientific question but lacks some clarity, completeness, or alignment with the given machine learning scenario.
* The question is appropriate for the type of machine learning analysis needed (classification or regression).
* Some details on goals or context are missing or underdeveloped.
 |
| 2 - Needs Improvement | * Attempts to define a scientific question but is vague, lacks key contextual elements, or does not sufficiently align with the goals and constraints of the analysis.
* The question is not appropriate for the type of machine learning analysis needed (classification or regression). Requires significant refinement.
 |
| 1 - Minimal | * Provides an incomplete or unclear scientific question that is mostly irrelevant to the analysis.
* The question is not appropriate for the type of machine learning analysis needed (classification or regression). Requires significant refinement.
 |
| 0 - No Response | * No response or entirely off-topic answer.
 |

| Part 2: Data HandlingExercise 2bDescribe your exploratory data analysis of any target and input features of note. Include the following:* Do variables follow diurnal or annual patterns generally as expected?
* Do the variables have the expected ranges of values? Do any variables appear to include major outliers?
* Which stations appear to be most correlated to the variables at Mt Mitchell?
* Include any important plots. Limit yourself to 5.
 |
| --- |
| **Score** | **Criteria** |
| 5 - Excellent | * Identifies whether variables follow diurnal and annual patterns as expected and explains any deviations.
* Evaluates if variables fall within expected ranges and identifies major outliers with justification.
* Analyzes correlations between stations and Mt. Mitchell, identifying the most correlated stations.
* Supports analysis with relevant plots and explains their significance.
 |
| 4 - Proficient | * Addresses most supporting questions with clear reasoning.
* Identifies diurnal/annual patterns, expected value ranges, and major outliers, though some discussion may lack depth or explain deviations
* Discusses station correlations but may not fully justify conclusions.
* Includes relevant plots but may not thoroughly explain all of them.
 |
| 3 - Satisfactory | * Covers some supporting questions but lacks depth or specificity in responses.
* Identifies diurnal and annual patterns but with limited explanation.
* Discusses expected value ranges but may overlook key outliers.
* Addresses station correlations but lacks strong supporting evidence.
* Includes some plots but does not clearly explain their significance.
 |
| 2 - Needs Improvement | * Addresses only a few supporting questions.
* Limited or unclear discussion of diurnal/annual patterns.
* Little to no evaluation of expected ranges or outliers.
* Mentions station correlations but with weak or no supporting analysis.
* Few or no relevant plots included.
 |
| 1 - Minimal | * Provides minimal or superficial responses, failing to address most supporting questions.
* Does not adequately analyze diurnal/annual patterns, expected value ranges, or outliers.
* No meaningful discussion of station correlations.
* Little to no supporting plots.
 |
| 0 - No Response | * No response or entirely off-topic answer.
 |

| Exercise 2cInput your data splitting strategy below. |
| --- |
| **Score** | **Criteria** |
| 5 - Excellent | * Training data between 60-80%
* Validation data between 10-20%
* Testing data between 10-20%
 |
| 1 - Minimal | * Percentages are present, but are not within the provided bounds for any of the three categories.
 |
| 0 - No Response | * No response or entirely off-topic answer.
 |

| Part 3: Model DevelopmentExercise 3ePaste evaluation resultsDescribe the results of your initial model validation. Include the following:* Which variables have favorable evaluation metrics? Which variables don’t perform as well?
* How do you interpret these statistics in the context of the physical world?
* What changes will you make to try to improve these statistics in the next iteration?
 |
| --- |
| **Score** | **Criteria** |
| 5 - Excellent | * Includes evaluation results.
* Provides a well-reasoned interpretation of validation metrics (RMSE and R²), and correctly identifies variables with favorable evaluation metrics and those that do not.
* Explains results (and unexpected results) in the context of the underlying science.
* Demonstrates a strong understanding of how input feature selection impacts results.
* Suggests potential improvements for subsequent trials.
 |
| 4 - Proficient | * Includes evaluation results.
* Provides a mostly accurate interpretation of validation metrics, correctly identifying most favorable and unfavorable variables.
* Discusses strengths and weaknesses of the model but may not fully explain unexpected results.
* Demonstrates a solid understanding of input feature selection but may not explore all implications.
* Suggests reasonable improvements, though some may lack depth.
 |
| 3 - Satisfactory | * Includes evaluation results.
* Provides a basic interpretation of validation metrics but may miss some key details in identifying which variables perform well or poorly.
* Mentions strengths and weaknesses of the model but does not fully explore their causes.
* Some recognition of input feature selection's impact, but explanation is limited.
* Suggests general improvements but lacks clear justification.
 |
| 2 - Needs Improvement | * Includes evaluation results.
* Attempts to interpret validation metrics but contains inaccuracies or lacks depth.
* Identifies some favorable or unfavorable variables but does not provide strong reasoning.
* Limited discussion of model strengths, weaknesses, or unexpected results.
* Little understanding of how feature selection affects results.
* Suggestions for improvement are vague or missing.
 |
| 1 - Minimal | * Includes evaluation results.
* Provides a superficial or incomplete response.
* Fails to correctly interpret validation metrics or identify which variables perform well or poorly.
* No meaningful discussion of model strengths, weaknesses, or unexpected results.
* Does not address the impact of feature selection.
* No actionable suggestions for improvement.
 |
| 0 - No Response | * No response or entirely off-topic answer.
 |

| Exercise 3fPaste the full output of each of your validation trials, one per box.  |
| --- |
| **Score** | **Criteria** |
| 5 - Excellent | * At least three unique additional trials are present
* Trials include more than one algorithm
* Trials use a variety input stations
 |
| 3 - Satisfactory | * Fewer than three unique additional trials are present
 |
| 0 - No Response | * No response or entirely off-topic answer.
 |

| Exercise 3hDescribe how your testing metrics compare to your validation metrics. Include the following:* Which environmental variables had the best evaluation metrics? List some physical scientific reasons why this may be the case.
* Is this model ready for use in the real world? Why or Why not?
* What other possible changes could further improve this model?
 |
| --- |
| **Score** | **Criteria** |
| 5 - Excellent | * Accurately identifies the environmental variables that had the best evaluation metrics.
* Suggests reasonable physical scientific reasons behind differences in variable performance.
* Thoughtfully assesses whether the model is ready for real-world use, providing strong justification within the context of the scientific issue at hand.
* Suggests concrete, scientifically valid improvements for future iterations.
 |
| 4 - Proficient | * Identifies most environmental variables with the best evaluation metrics, though minor inaccuracies may be present.
* Provides reasonable scientific explanations for variable performance but may lack depth in some areas.
* Assesses real-world readiness with justification, though reasoning may not be fully developed.
* Suggests meaningful improvements, but they may not be fully explained.
 |
| 3 - Satisfactory | * Identifies some of the best-performing variables but with partial accuracy or missing details.
* Offers a basic scientific explanation for variable performance but lacks depth.
* Provides a general assessment of real-world readiness, though justification is weak or incomplete.
* Suggests potential improvements but with little scientific reasoning.
 |
| 2 - Needs Improvement | * Attempts to identify well-performing variables but with significant inaccuracies or missing key metrics.
* Provides little or unclear scientific reasoning behind variable performance.
* Minimal discussion of real-world readiness, with weak or unsupported justification.
* Suggestions for improvement are vague or not scientifically valid.
 |
| 1 - Minimal | * Provides an incomplete or superficial response.
* Fails to correctly identify well-performing variables or provide scientific reasoning.
* Offers little to no discussion on real-world readiness.
* No meaningful suggestions for improvement.
 |
| 0 - No Response | * No response or entirely off-topic answer.
 |