Test Results and Interview Guide

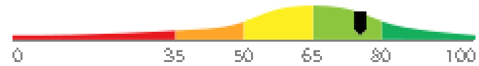
Candidate:	**Elizabeth Wantsajob**
Assessment:	Machine Learning Concepts
Completed:	July 2, 2026
Prepared for:	Sara Maple Example Company

What's Included

- Overall Score
- Competency Summary Table
- Comparison Matrix
- Detailed Competency Results with Interview Guide

Important Note: The Machine Learning Concepts assessment measures key factors related to high performance and tenure in this job. Attribute types measured vary by test, but can include cognitive ability, skills, knowledge, personality characteristics, emotional intelligence, and past behavioral history. This report includes a one page summary, followed by detailed results with an embedded interview guide. Note that these results should always be used as a part of a balanced candidate selection process that includes independent evaluation steps, such as interviews and reference checks.

Overall

Candidate	Score	Interpretation
Elizabeth Wantsajob	75	

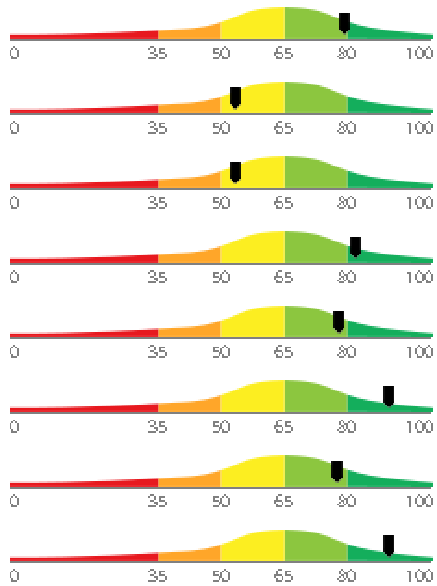
beth.wantsajob@gmail.com
Machine Learning Concepts
July 2, 2026

The candidate demonstrates a solid and competent understanding of Machine Learning concepts, including supervised and unsupervised learning, common algorithms, evaluation metrics, and data preprocessing techniques. Minor gaps may exist in advanced areas such as deep learning frameworks or complex hyperparameter tuning strategies, but this individual is well-positioned to contribute effectively in an entry-level to mid-level capacity.

Key
- Candidate Score
- Higher Risk
- Lower Risk

Competency Summary

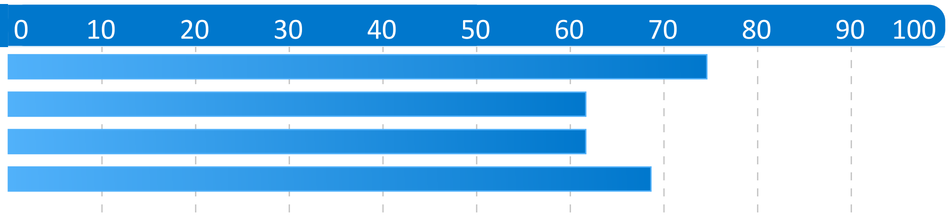
Competency	Score	Interpretation
Skills/Knowledge (relates to immediate readiness)		
Data Preprocessing and Feature Engineering	79	
Model Evaluation and Metrics (Free Text Responses)	53	
Supervised Learning Algorithms (Free Text Responses)	53	
Model Evaluation and Metrics	82	
Overfitting, Underfitting, and Model Generalization	78	
Supervised Learning Algorithms	90	
Training, Validation, and Test Set Splits	77	
Unsupervised Learning and Clustering	90	



Comparison

Percentile scores indicate how the candidate compares to other test-takers within various groups. The candidate scored equal to or better than the fraction of test-takers indicated by the percentile.

Test-Taker Group	Percentile
Global	75th
North America	62nd
United States	62nd
Example Company	69th



Artificial Intelligence (AI) Generated Scores

This table includes one or more scores derived from a large language model AI query. AI-derived scores are non-deterministic. That is, they are not precisely repeatable. Therefore, these scores should always be treated as supplementary information and should never be used exclusively or compared to hard cutoff values.

Estimated Value	Score	Confidence	Interpretation
Knowledge, Skills, and Abilities Summary	-	-	Summary Points (AI): • (Generic Text for Sample Report) Strong performer in Drag and Drop Files tasks, indicating comfort with file management and basic computer interactions. • Demonstrates solid numerical accuracy in Recognizing and Confirming Numbers, a valuable asset in detail-oriented roles. • Moderate overall performance in Analytical Thinking and Attention to Detail, with adequate grammar skills but room for improvement. • Struggles with Reading and Analyzing Problems, which may limit effectiveness in roles requiring critical reading and complex problem-solving. • Lowest performance in Navigating Between Screens, suggesting difficulty with multi-screen software workflows that could impact productivity in computer-intensive roles. Narrative (AI): Elizabeth Wantsajob demonstrates a mixed profile of knowledge, skills, and abilities across the assessed competencies. Elizabeth shows a strong aptitude in Drag and Drop Files, performing well on this technical task and suggesting she is comfortable with this type of computer interaction. This is a notable strength that would translate well into roles requiring file management and basic computer navigation tasks. In the area of Analytical Thinking and Attention to Detail, Elizabeth performs at a moderate level. She demonstrates solid ability in Recognizing and Confirming Numbers, which suggests she is careful and accurate when working with numerical data — a valuable skill in detail-oriented work environments. Her Grammar performance is adequate but leaves room for improvement, indicating she may occasionally make written communication errors. Her weakest area within this competency is Reading and Analyzing Problems, where she struggled to consistently interpret and work through written problem scenarios. This may impact her effectiveness in roles that require critical reading, written comprehension, or complex problem-solving. Elizabeth's most significant area for development is Navigating Between Screens, where she scored considerably lower than the other competencies. This suggests she may have difficulty efficiently moving through software interfaces or multi-screen workflows, which could slow productivity in roles that rely heavily on navigating computer applications or data entry systems. Overall, Elizabeth brings some useful technical strengths, particularly in file management and numerical accuracy, but would benefit from targeted development in software navigation and analytical problem-solving to be fully effective in roles that demand these skills. Computed on: April 2, 2026, 11:09:49PM EDT

Detail

Candidate:	Elizabeth Wantsajob, beth.wantsajob@gmail.com
Assessment:	Machine Learning Concepts
Authorized:	July 2, 2026, by Sara Maple, Example Company, qamailsaram.mike@hravatar.com
Started:	July 2, 2026, 5:21:10PM EDT
Completed:	July 2, 2026, 5:21:10PM EDT
Overall Score:	75

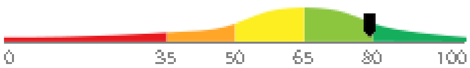
Knowledge and Skills Detail

This section contains a list of job-related knowledge areas and skills that have been evaluated. Low scores in these areas often indicate that additional learning may be required before top performance can be achieved.

Detail | Interview Guide

Data Preprocessing and Feature Engineering

Score: 79



Description:

Covers the practical steps needed to prepare raw data for use in machine learning models, including handling missing values, encoding categorical variables, normalizing or scaling numeric features, and creating or selecting the most useful features. These tasks are performed in nearly every real-world machine learning project.

Interpretation:

Candidate should achieve above average job performance in this area with little or no training.

The candidate exhibits a solid and largely consistent understanding of data preprocessing and feature engineering, and is capable of independently executing most standard tasks required to prepare data for machine learning models. Minor gaps in knowledge may exist in advanced feature selection or engineering techniques, but overall competence is well-established.

Describe a feature engineering step you have applied or would apply to improve model performance, and explain why it would help.

1	2	3	4	5
Cannot describe a concrete feature engineering step or explain its impact on model performance.		Describes a basic step such as encoding or scaling but provides limited explanation of why it improves performance.		Describes a meaningful transformation such as interaction terms, binning, or log transformation with a clear, reasoned explanation of its impact.

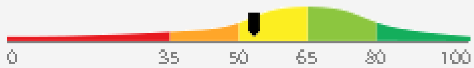
Why is it important to normalize or scale numeric features before training certain machine learning models?

1	2	3	4	5
Cannot explain why scaling matters or which types of models are affected by unscaled features.		States that scaling helps model performance but cannot explain the underlying reason or specify which algorithms are affected.		Explains that distance- or gradient-based models are sensitive to feature scale and gives examples such as k-nearest neighbors or gradient descent-based models.

Detail

Model Evaluation and Metrics (Free Text Responses)

Score: 53



Description:

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

Interpretation:

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

Overall AI Score:	60.0
High words per minute detected while composing one or more essays:	27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15.
AI Confidence Level:	80
Argument Strength (AI):	70.0
Clarity and Coherence (AI):	80.0
Match with Ideal Response (AI):	60.0
Other Errors per 100 Words:	0.0
Spelling errors per 100 words:	0.0

Please see below to view the essay submitted.

Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

1	2	3	4	5
Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases.		Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions.		Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track.

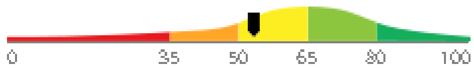
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

1	2	3	4	5
Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines.		Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout.		Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases.

Detail

Supervised Learning Algorithms (Free Text Responses)

Score: 53



Description:

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

Interpretation:

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

Overall AI Score:	60.0
High words per minute detected while composing one or more essays:	27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15.
AI Confidence Level:	80
Argument Strength (AI):	70.0
Clarity and Coherence (AI):	80.0
Match with Ideal Response (AI):	60.0
Other Errors per 100 Words:	0.0
Spelling errors per 100 words:	0.0

Please see below to view the essay submitted.

Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

1	2	3	4	5
Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases.		Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions.		Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track.

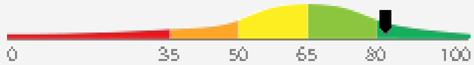
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

1	2	3	4	5
Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines.		Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout.		Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases.

Detail	Interview Guide

Model Evaluation and Metrics

Score: 82



Description:

Focuses on how to measure and interpret model performance using metrics such as accuracy, precision, recall, F1 score, and RMSE. Knowing which metric to use for a given business problem and how to interpret results is critical for determining whether a model is ready for real-world use.

Interpretation:

Candidate should achieve superior job performance in this area with little or no training.

The candidate exhibits an advanced and comprehensive mastery of model evaluation and metrics within machine learning. They can confidently select, apply, and interpret a wide range of performance metrics in the context of diverse business problems, and are well-equipped to make informed, data-driven decisions regarding model readiness for production use.

A business stakeholder wants to minimize false negatives in a fraud detection model. Which evaluation metric would you prioritize and why?

1	2	3	4	5
Cannot connect the business goal to a specific metric or confuses recall and precision.		Identifies recall as the relevant metric but provides limited reasoning tied to the business context.		Correctly prioritizes recall, explains the cost of false negatives in fraud detection, and discusses potential trade-offs with precision.

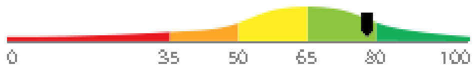
What is the difference between accuracy and precision, and why might accuracy alone be misleading for some problems?

1	2	3	4	5
Cannot distinguish between accuracy and precision or explain why accuracy can be misleading.		Defines both terms correctly but struggles to explain when accuracy is misleading or why.		Clearly defines both metrics and explains class imbalance as a scenario where accuracy misleads, with a concrete example.

Detail

Overfitting, Underfitting, and Model Generalization

Score: 78



Description:

Addresses the common challenge of building models that perform well on new, unseen data. Covers the concepts of bias and variance, and practical techniques such as cross-validation and regularization that help prevent a model from being too closely fit to training data or too simple to capture meaningful patterns.

Interpretation:

Candidate should achieve above average job performance in this area with little or no training.

The candidate exhibits a solid and well-rounded understanding of model generalization, including the bias-variance tradeoff and practical techniques such as cross-validation and regularization. Minor gaps in knowledge or application may exist, but overall competence in this area is evident.

Interview Guide

Walk me through how you would use cross-validation to evaluate a model and why it gives a better estimate of performance than a single train-test split.

1	2	3	4	5
Cannot explain the cross-validation process or why it is preferable to a single split.		Describes the basic k-fold process but does not clearly explain the advantage over a single split.		Accurately describes k-fold cross-validation, explains variance reduction in performance estimates, and notes its value for small datasets.

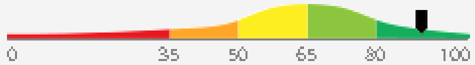
What does it mean for a model to overfit, and how would you know if your model is overfitting?

1	2	3	4	5
Cannot define overfitting or describe any signs that a model is overfitting.		Defines overfitting correctly but can only vaguely describe how to detect it.		Defines overfitting clearly, explains the train-vs-validation performance gap, and mentions specific detection methods like cross-validation.

Detail

Supervised Learning Algorithms

Score: 90



Description:

Covers the most commonly used machine learning algorithms for classification and regression tasks, including linear regression, logistic regression, decision trees, and random forests. Understanding these algorithms—how they work, when to use them, and their strengths and weaknesses—is foundational for building predictive models in business applications.

Interpretation:

Candidate should achieve superior job performance in this area with little or no training.

The candidate exhibits a strong and comprehensive mastery of machine learning concepts, including algorithm selection, model optimization, bias-variance tradeoffs, neural networks, and proficiency with industry-standard libraries and frameworks. They are well-equipped to independently design, develop, evaluate, and refine machine learning models for a wide range of business applications with a high degree of competence.

Interview Guide

When would you choose a random forest model over a single decision tree, and what trade-offs would you consider?

1	2	3	4	5
Cannot articulate differences or trade-offs between random forest and a single decision tree.		Mentions reduced overfitting or improved accuracy but lacks depth on trade-offs like interpretability or computation.		Clearly explains ensemble benefits, variance reduction, and thoughtfully discusses trade-offs such as interpretability and training cost.

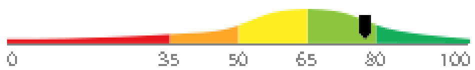
Can you explain in simple terms what a decision tree does and give an example of a business problem where you might use one?

1	2	3	4	5
Cannot explain how a decision tree works or provide a relevant business example.		Gives a basic explanation of decision trees and a general business example with limited detail.		Clearly explains decision tree logic, splitting criteria, and provides a specific, well-reasoned business use case.

Detail

Training, Validation, and Test Set Splits

Score: 77



Description:

Covers how to properly divide data into training, validation, and test sets to support reliable model development and evaluation. Understanding why each split is needed and how to use them correctly prevents common mistakes like data leakage and overly optimistic performance estimates.

Interpretation:

Candidate should achieve above average job performance in this area with little or no training.

The candidate exhibits a solid understanding of how to properly partition data into training, validation, and test sets to support reliable model development and evaluation. They are generally capable of applying correct splitting strategies and are aware of common pitfalls such as data leakage, though some advanced nuances may still require development.

Interview Guide

What is data leakage, and how can it occur when splitting your dataset for model training and evaluation?

1	2	3	4	5
Cannot define data leakage or describe how it can occur during dataset splitting.		Defines data leakage at a high level but cannot give a specific example of how it happens during splitting.		Clearly defines data leakage, gives a concrete example such as scaling before splitting, and explains how it inflates performance metrics.

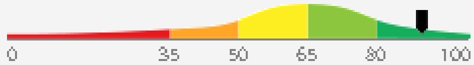
Can you explain the difference between a training set and a test set, and why we keep them separate?

1	2	3	4	5
Cannot explain the purpose of each set or why separation is necessary.		Correctly states the basic purpose of each set but does not explain the risk of using test data during training.		Clearly explains both sets' roles and articulates that using test data during training leads to overly optimistic and unreliable performance estimates.

Detail	Interview Guide

Unsupervised Learning and Clustering

Score: 90



Description:

Covers machine learning approaches used when labeled data is not available, with a focus on clustering techniques such as k-means. These methods are widely used in business for tasks like customer segmentation, anomaly detection, and pattern discovery in large datasets.

Interpretation:

Candidate should achieve superior job performance in this area with little or no training.

The candidate exhibits an advanced and comprehensive mastery of unsupervised learning and clustering concepts, reflecting deep expertise in both theory and practical application. They are well-equipped to lead efforts involving complex tasks such as large-scale pattern discovery, anomaly detection, and customer segmentation using sophisticated machine learning approaches.

How does the k-means clustering algorithm work, and how would you decide on the right number of clusters for a business problem?

1	2	3	4	5
Cannot explain how k-means works or suggest any method for choosing the number of clusters.		Describes the basic k-means process but cannot explain a systematic method for choosing k.		Accurately explains k-means iteration and centroid assignment, and describes methods like the elbow method or silhouette score for selecting k.

What is the difference between supervised and unsupervised learning, and can you give a business example of each?

1	2	3	4	5
Cannot clearly distinguish between supervised and unsupervised learning or provide relevant examples.		Correctly distinguishes the two types but provides only generic or vague business examples.		Clearly distinguishes the two approaches and provides specific, realistic business examples for each, such as churn prediction vs. customer segmentation.

Free Text Responses

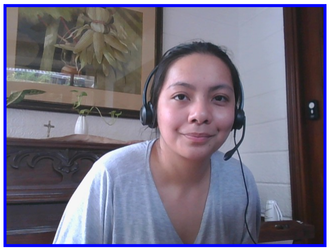
During the assessment, the candidate was asked to answer one or more questions using text, audio, video, or an uploaded text file. Their responses are included below for review.

Question or Task	Response
After an AI product is deployed, what is model monitoring and why is it a necessary part of the product lifecycle?	Model monitoring is a technique for ensuring that the model does not wander or become overtrained after an extended period of repeated queries that have the same or similar prompts. This is very important for preventing hallucinization. It's also a key aspect of any guardrails strategy. **Comments (AI):** The answer is clear and coherent but lacks depth in explaining the importance of model monitoring. The phrase 'hallucinization' is not commonly used in this context and may confuse readers. The answer could be improved by providing more specific examples of model performance metrics and how they are tracked. The argument strength is moderate as it does not fully explain why model monitoring is necessary in the product lifecycle. **Misspelled Words:** guardrails (1), hallucinization (1)

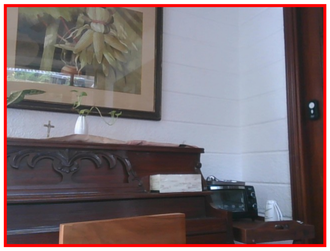
Identity Confirmation Photos

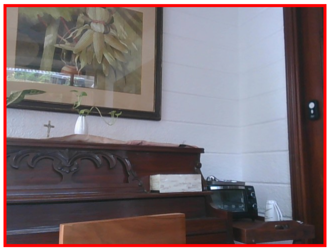
The following photos of the candidate and any identification were uploaded during the assessment session.

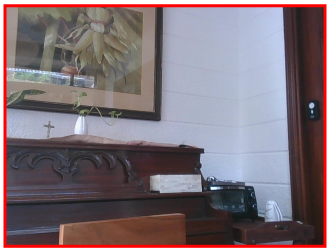
Photo Analysis Results	
- Risk:	Medium risk of cheating based on image inconsistencies
- Percent match among processed faces	100%
- Total images processed	17
- Total images with valid faces	14 (82%)
- Total pairs of faces compared	13
- Pairs in which faces matched	13 (100%)


Pre/Post-Test Photo

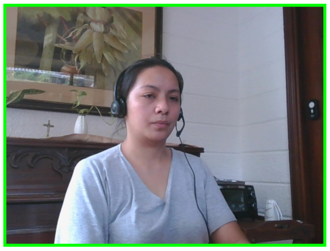

ID Photo

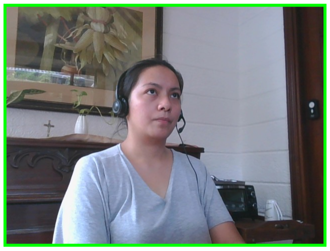

In-Test Error Detected (No Face Detected)

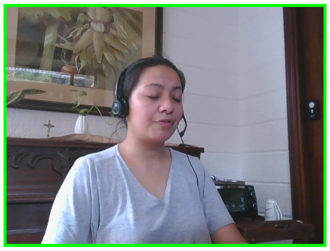

In-Test Error Detected (No Face Detected)

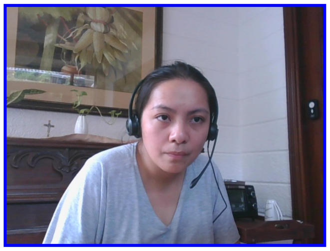

In-Test Error Detected (No Face Detected)


In-Test Photo


In-Test Photo


In-Test Photo


In-Test Photo


Pre/Post-Test Photo

Resume or CV

Summary Updated on

Motivated career professional with extensive experience in office administration and management. Proven track record of improving efficiency, reducing costs, and enhancing office operations through strategic initiatives and technology implementation.

Objective

I am seeking a role where I can use my many skills and my exceptional judgment and empathy for customers to make a difference to a growing company.

Education

- Associate of Applied Science in Office Administration, Portland Community College, 2020

Experience

- General Office Clerk, Paramount Office Management, 09/2023 – Present
- Administrative Assistant, Global Enterprises Inc., 04/2021 – 08/2023
- Administrative Assistant, Innovative Business Solutions Ltd., 07/2019 – 03/2021

Other Qualifications

- Microsoft Office Specialist (MOS) Certification
- Certified Administrative Professional (CAP)
- International Association of Administrative Professionals (IAAP) Certification

Report Preparation Notes

- Hiring decisions should never be based on a single source of information. The most effective use of this assessment report is as a part of a multi-faceted program of candidate evaluation that includes resume review, interviews, and reference checks.
- Overall vs Percentiles Scores: The overall score reflects the success in the test, based on the mean (average) and standard deviation of the test scores. The percentile score reflects the percentage of test-takers who scored equal or below this overall score. We recommend you use the Overall Score as your primary evaluation criteria. However, percentile scores can often be useful in comparing specific candidates against one another and with a group, such as for test takers in a certain organization or within a certain account.
- Note that comparison information is calculated based on completed instances of this assessment at that time the assessment is scored. As additional instances are completed, the comparative data may change. You can always update a report to the current values by clicking on 'Recalculate Percentiles' within the online results viewing pages at www.hravatar.com.
- Most competency scores are norm-based, which means that they can be interpreted in terms of their distance from the average or mean score. For all scales, a score equal to the mean receives a score of 65 and scores above and below this value are set so that a score change of 15 equals one standard deviation.
- For linear competencies, higher is better across the entire scale. For these scales a score between 65 and 80 (light green) represents 0 to 1 standard deviation above the mean and a score above 80 (dark green) represents more than one standard deviation above the mean. Similarly, a score of 50 - 65 (yellow) represents 0 to 1 standard deviation below the mean, while a score of 35 - 50 (orange) equates to 1 to 2 standard deviations below the mean, and a score below 35 represents more than 2 standard deviations below the mean.
- Sim ID: 20889-1, Key: 0-0, Rpt: 68, Prd: 9709, Created: 2026-07-02 17:21 EDT
- UA: Mozilla/5.0 (Windows NT 6.3; Trident/7.0; Touch; rv:11.0) like Gecko

Score Calculation Detail

The following table provides a summary of how the overall score was calculated from each of the individual competency scores. First, all competency scores are calculated on a scale of 0-100. Note that some competencies use their color category rather than their actual numeric score in the overall calculation. For these, a standard score associated with the assigned color category is used in the overall score calculation rather than the actual numeric score. This is reflected in the ''Score Value Used'' column. Next, a weighted average of scores is computed using individual competency weights, typically set using job analysis data provided by the US Government Occupational Information Network (O*Net).

Competency	Score	How applied to overall	Score Value Used	Weight (%)
Data Preprocessing and Feature Engineering	79.5711	Numeric Score	79.5711	12.5000
Model Evaluation and Metrics	82.2456	Numeric Score	82.2456	12.5000
Model Evaluation and Metrics (Free Text Responses)	53.8624	Numeric Score	53.8624	12.5000
Overfitting, Underfitting, and Model Generalization	78.4296	Numeric Score	78.4296	12.5000
Supervised Learning Algorithms	90.0705	Numeric Score	90.0705	12.5000
Supervised Learning Algorithms (Free Text Responses)	53.8624	Numeric Score	53.8624	12.5000
Training, Validation, and Test Set Splits	77.7120	Numeric Score	77.7120	12.5000
Unsupervised Learning and Clustering	90.2399	Numeric Score	90.2399	12.5000
Weighted Average:				75.7492
Final Overall Score:				75

Notes

(This area is intentionally blank - it's reserved as space for your notes.)