

Machine Learning Engineer

FF08 Data, analytics and AI · ISCO-08 2512 · CBS 0814 Software and application developers · seniority senior · Also known as: MLOps Engineer, ML Platform Engineer

Takes models to production and keeps them working by automating and managing training, deployment and monitoring of model performance.

The machine learning engineer turns a model into a service that delivers predictions daily within agreed response times and cost. Unlike the data scientist, research is not the end product, because this role owns model behaviour in production and the fallback when it goes wrong. Selection tests model knowledge, while the role depends on software engineering and on detecting in time that reality has shifted and the model quietly returns wrong outcomes.

Key figures for this profile

28 skills · 3 competencies to measure · 6 knockout skills · centre of gravity: Digital, data and AI 73%, Cognitive ability and problem solving 15%, Communication and influence 5%

1 Competencies to measure

These are the behavioural and ability constructs in this profile. Each one shows which hrmforce questionnaire measures it.

Competencies to measure	Target level	hrmforce instrument	Assessment method	Behavioural anchor at target level
Numerical reasoning Knockout	■■■■■ N5	Ability Scan	Cognitive ability test	Sees in complex number series which assumptions determine the outcome and sets the calculation logic that others in the organisation follow.

Competencies to measure	Target level	hrmforce instrument	Assessment method	Behavioural anchor at target level
Curiosity and exploring Competency (50-framework): Learning Ability		Big Fifty, 15PF, Motivation	Personality questionnaire	Gathers information on new methods outside the own team and brings findings into the work meeting unprompted.
Dealing with ambiguity Competency (50-framework): Flexibility		Big Fifty, Mental Resilience (4C), Competency Check	Situational Judgement Test	Formulates explicit assumptions when information is missing, proceeds on that basis and revises the approach once new data arrives.

2 Full skill profile

Every skill in this profile, sorted by weight. The behavioural anchor under each skill belongs to the target level.

T	Skill	Target level	Weight	Knockout	Assessment method	hrmforce instrument
A	Numerical reasoning Sees in complex number series which assumptions determine the outcome and sets the calculation logic that others in the organisation follow.	■■■■■ N5	5	yes	Cognitive ability test	Ability Scan
V	Data analysis Sets the organisation's analysis standards, reviews the analyses of others and establishes new analysis methods in the profession.	■■■■■ N5	5	yes	Work sample test	Ability Scan, Work sample test
K	Basic statistical interpretation Reviews the statistical basis of organisation wide decisions and trains others in the correct interpretation of results.	■■■■■ N5	5	yes	Knowledge test	Ability Scan, Knowledge test (client-specific)
K	Machine learning fundamentals Sets requirements for model development in the organisation and reviews third party models on design and evidence.	■■■■■ N5	5	yes	Knowledge test	Knowledge test (client-specific)

T	Skill	Target level	Weight	Knockout	Assessment method	hrmforce instrument
V	Programming Independently builds a function or module from the requirement, chooses suitable structures and keeps the code readable and tested.	 N4	5	yes	Work sample test	Work sample test, Knowledge test (client-specific)
V	Critical thinking and source evaluation Sets the assessment criteria for sources and claims, including for texts made by artificial intelligence, and enforces them organisation wide.	 N5	4	no	Work sample test	Competency Check, Work sample test
V	Data visualisation Sets the organisation's visualisation guidelines and judges whether reports by others represent the data honestly.	 N5	4	no	Work sample test	Work sample test, Portfolio
V	Assessing data quality Establishes the data quality framework for the entire organisation and decides which sources count as reliable.	 N5	4	yes	Work sample test	Work sample test, Knowledge test (client-specific)
V	Data modelling Sets the organisation's modelling standards and core data model and decides on deviations from them.	 N5	4	no	Work sample test	Knowledge test (client-specific), Work sample test

T	Skill	Target level	Weight	Knockout	Assessment method	hrmforce instrument
V	Software design and architecture Sets the organisation's architecture principles and decides on designs that affect multiple systems.	■■■■■ N5	4	no	Work sample test	Work sample test, Portfolio
T	Version control and collaboration workflow Determines the way of working for version control and code review in the organisation and enforces it across teams.	■■■■■ N5	4	no	Work sample test	Work sample test
V	Release management and deployment Sets the organisation's release policy, decides on rollout windows and structurally improves the lead time of changes.	■■■■■ N5	4	no	Work sample test	Knowledge test (client-specific), Work sample test
V	Using cloud platforms Determines the organisation's cloud strategy and decides which workloads are placed where.	■■■■■ N5	3	no	Knowledge test	Knowledge test (client-specific)
V	Systems thinking Maps the chain around a bottleneck, points out the feedback loop and predicts the effect of an intervention.	■■■■■ N4	3	no	Work sample test	Competency Check, Work sample test
V	Methodical working Chooses a suitable method for a new task, records the steps and sticks to them under time pressure.	■■■■■ N4	3	no	On-the-job observation	Competency Check, Work sample test

T	Skill	Target level	Weight	Knockout	Assessment method	hrmforce instrument
V	Advisory skill Reframes the question behind the question, offers two supported options with consequences and agrees a decision moment.	 N4	3	no	Assessment centre	Competency Check, Structured interview
G	Curiosity and exploring Gathers information on new methods outside the own team and brings findings into the work meeting unprompted.	 N4	3	no	Personality questionnaire	Big Fifty, 15PF, Motivation
V	Testing hypotheses Turns a question into a testable hypothesis, chooses the test that fits the data and states the limitations of the conclusion.	 N4	3	no	Work sample test	Ability Scan, Work sample test
V	Data management and storage Independently manages data collections, plans maintenance and recovery actions and tests whether a restore actually works.	 N4	3	no	Work sample test	Knowledge test (client-specific), Work sample test
V	Data governance and ownership Assigns ownership for a data domain, has definitions agreed and records decision making about data use traceably.	 N4	3	no	Structured interview	Structured interview, Portfolio

T	Skill	Target level	Weight	Knockout	Assessment method	hrmforce instrument
K	GDPR and privacy in practice Judges in own processes whether a legal basis exists, applies data minimisation and handles requests from data subjects correctly.		3	no	Knowledge test	Knowledge test (client-specific)
V	Software testing and quality assurance Designs test cases including edge cases, automates repeatable tests and advises on release based on the findings.		3	no	Work sample test	Work sample test, Knowledge test (client-specific)
V	Selecting and implementing AI applications Compares applications on requirements, risk and cost, runs a pilot and guides colleagues through adoption.		3	no	Structured interview	Structured interview, Portfolio
V	AI governance and responsible use Tests intended AI use against policy and legislation, records the trade off and arranges human oversight of decisions.		3	no	Knowledge test	Knowledge test (client-specific), Structured interview
V	Recognising AI risk and bias Examines outcomes per group or situation, identifies where the system falls short and proposes measures or extra checks.		3	no	Work sample test	Knowledge test (client-specific), Work sample test

T	Skill	Target level	Weight	Knockout	Assessment method	hrmforce instrument
V	multidisciplinary collaboration Maps the issue with specialists from different disciplines and formulates a joint approach with a clear division of roles.	 N4	2	no	Assessment centre	360 Feedback, Competency Check
V	Visual communication and data storytelling Creates a chart in an existing template, adds a title and names the unit shown on the axis.	 N2	2	no	Work sample test	Work sample test, Portfolio
G	Dealing with ambiguity Formulates explicit assumptions when information is missing, proceeds on that basis and revises the approach once new data arrives.	 N4	1	no	Situational Judgement Test	Big Fifty, Mental Resilience (4C), Competency Check

3 Levels per seniority

The same profile with the job family seniority adjustment applied. Target levels are clamped to 1 to 5.

Skill	medior	senior
Numerical reasoning	N4	N5
Data analysis	N4	N5
Basic statistical interpretation	N4	N5
Machine learning fundamentals	N5	N5
Programming	N4	N4
Critical thinking and source evaluation	N4	N5
Data visualisation	N4	N5
Assessing data quality	N4	N5
Data modelling	N4	N5
Software design and architecture	N4	N5
Version control and collaboration workflow	N4	N5
Release management and deployment	N4	N5
Using cloud platforms	N4	N5
Systems thinking	N3	N4
Methodical working	N3	N4
Advisory skill	N3	N4
Curiosity and exploring	N3	N4
Testing hypotheses	N3	N4
Data management and storage	N3	N4
Data governance and ownership	N3	N4
GDPR and privacy in practice	N3	N4
Software testing and quality assurance	N3	N4
Selecting and implementing AI applications	N3	N4

Skill	medior	senior
AI governance and responsible use	N3	N4
Recognising AI risk and bias	N3	N4
multidisciplinary collaboration	N3	N4
Visual communication and data storytelling	N2	N2
Dealing with ambiguity	N3	N4

4 The proficiency scale

Level	Label	Autonomy	Complexity	Scope	Depth of knowledge
N1	Guided	works under supervision and follows instruction	routine, one variable at a time	own task	basic understanding of terminology
N2	Applying	works independently within set frameworks	familiar situations with limited variation	own role	working knowledge, applies standard approach
N3	Proficient	sets own approach and seeks input proactively	several variables, some ambiguity	own team or process	thorough knowledge, weighs alternatives
N4	Advanced	directs others and decides on the approach	complex, with conflicting interests	several teams or a department	in-depth knowledge, advises others
N5	Leading	sets the standard and the policy	strategic, under high uncertainty	organisation, value chain or profession	authority, advances the profession

5 Behavioural anchors per skill

Anchors sit at N1, N3 and N5. N2 and N4 are deliberately not anchored: raters interpolate between them, following the O*NET convention.

Skill	N1	N3	N5
Numerical reasoning Works with numbers, ratios and tables and draws conclusions from them about quantities, trends and relationships in numerical material.	Reads a simple table, names the highest and lowest value and calculates a percentage using a calculator.	Combines two data sources, calculates the development over time and explains which calculation step leads to which result.	Sees in complex number series which assumptions determine the outcome and sets the calculation logic that others in the organisation follow.
Data analysis Prepares and analyses data sets to find patterns, relationships and outliers and connects the outcomes to the original question.	Carries out a predefined analysis step on a supplied data set and records the outcome in the agreed format.	Selects analysis methods for an open question, turns raw data into a working file and supports every conclusion with figures.	Sets the organisation's analysis standards, reviews the analyses of others and establishes new analysis methods in the profession.

Skill	N1	N3	N5
Basic statistical interpretation Knows and interprets basic statistical concepts such as spread, correlation, confidence interval and significance and states what an outcome does and does not show.	Reads mean, spread and counts correctly from a table and checks what a percentage exactly represents.	Interprets correlations, confidence intervals and significance in a report and points out where cause and effect are not established.	Reviews the statistical basis of organisation wide decisions and trains others in the correct interpretation of results.
Machine learning fundamentals Knows the main forms of machine learning, the role of training data and concepts such as overfitting, validation and model performance in plain language.	States the difference between supervised and unsupervised learning and which data a model needs to learn.	Explains how a model is trained and validated and judges whether reported performance figures are credible.	Sets requirements for model development in the organisation and reviews third party models on design and evidence.
Programming Writes, reads and improves working program code in at least one language, splits a problem into steps and tests own solution.	Adjusts existing code as directed, runs it and reports error messages that cannot be resolved alone.	Independently builds a function or module from the requirement, chooses suitable structures and keeps the code readable and tested.	Sets programming standards and language choices for the organisation and reviews others' code for quality and maintainability.
Critical thinking and source evaluation Tests claims for logic, evidence and the interest of the source and thereby distinguishes facts from opinions and assumptions.	Asks what a claim is based on and checks whether the source is named and can be found.	Detects fallacies and selective evidence in reports and states which conclusion the data do support.	Sets the assessment criteria for sources and claims, including for texts made by artificial intelligence, and enforces them organisation wide.
Data visualisation Turns data into charts and visual displays that show the message correctly, chooses suitable chart types and avoids misleading presentation.	Creates a chart following an example, completes axes and labels and checks whether the figures are correct.	Chooses the chart type per message, removes redundant elements and explains why an alternative display would mislead.	Sets the organisation's visualisation guidelines and judges whether reports by others represent the data honestly.

Skill	N1	N3	N5
Assessing data quality Checks data sets for completeness, accuracy, duplicates and timeliness, describes the defects found and determines whether the data is usable.	Works through a checklist on a supplied data set and reports missing or duplicate records to the person responsible.	Defines validation rules for a data source, quantifies the defects and advises whether the data is fit for use.	Establishes the data quality framework for the entire organisation and decides which sources count as reliable.
Data modelling Describes data and its relationships in a model with entities, attributes and keys that matches the work process.	Extends an existing data model with new fields following the applicable naming conventions and has the work reviewed.	Designs a logical data model for a defined domain, normalises where needed and discusses choices with users and developers.	Sets the organisation's modelling standards and core data model and decides on deviations from them.
Software design and architecture Designs the structure of an application or interface, distributes functionality across components and weighs choices on maintainability, performance and security.	Elaborates a supplied design into components and asks questions when the design turns out to be incomplete.	Produces a design for a defined application, records the main choices and alternatives and reviews it with colleagues.	Sets the organisation's architecture principles and decides on designs that affect multiple systems.
Version control and collaboration workflow Works with a version control system, uses branches and merge requests and keeps the change history traceable for the team.	Commits changes with a clear description in the designated branch according to the team agreement.	Manages branches and merges independently, resolves conflicts and reviews colleagues' changes on substance.	Determines the way of working for version control and code review in the organisation and enforces it across teams.
Release management and deployment Plans and carries out the rollout of new versions, checks operation before and after and ensures a rollback stays possible.	Executes the steps of a rollout plan and reports immediately when a step does not give the expected result.	Draws up a rollout plan with a fallback option, aligns with stakeholders and verifies after rollout that the service works.	Sets the organisation's release policy, decides on rollout windows and structurally improves the lead time of changes.

Skill	N1	N3	N5
Using cloud platforms Sets up and manages services in a cloud environment, arranges permissions and costs and chooses between available service types based on requirements.	Uses existing cloud services as instructed and reports unusual behaviour or unexpected costs to the administrator.	Independently configures cloud services for an application, arranges permissions and backups and keeps costs and usage in view.	Determines the organisation's cloud strategy and decides which workloads are placed where.
Systems thinking Describes how parts of a process or organisation influence each other, including feedback, delay and unintended effects.	Names who or what depends on the own task earlier and later in the process.	Maps the chain around a bottleneck, points out the feedback loop and predicts the effect of an intervention.	Models the interplay between departments, market and regulation and shifts the intervention to the point with the greatest leverage.
Methodical working Follows a fixed and traceable method for tasks, with steps, documentation and checks, so that others can repeat the work.	Completes the steps of a prescribed procedure in the right order and records the result.	Chooses a suitable method for a new task, records the steps and sticks to them under time pressure.	Writes the working methods for the field, has them tested in practice and revises them at fixed moments.
Advisory skill Sharpening the question of a client and delivering a well founded recommendation in a way that the receiver adopts and implements it.	Supplies requested information and explains a standard recommendation following the fixed structure used by the team.	Reframes the question behind the question, offers two supported options with consequences and agrees a decision moment.	Advises the board on issues with conflicting interests, challenges established assumptions and safeguards the quality of advice across the organisation.
Curiosity and exploring Asks questions about the unfamiliar, seeks out new topics and perspectives unprompted and explores them further than the work strictly requires.	Asks follow up questions during an explanation about a new topic and reads the background material provided.	Gathers information on new methods outside the own team and brings findings into the work meeting unprompted.	Explores developments outside the sector systematically, connects unexpected findings and puts new themes on the agenda of the organisation.

Skill	N1	N3	N5
Testing hypotheses Formulates a testable assumption, chooses a suitable test or comparison, carries it out and draws a conclusion with its uncertainty.	Carries out a provided test according to a step plan and records the outcome without adding own interpretation.	Turns a question into a testable hypothesis, chooses the test that fits the data and states the limitations of the conclusion.	Designs research set ups for the organisation and judges whether conclusions of others are supported by the data.
Data management and storage Manages storage, access and lifecycle of data collections, applies changes in a controlled way and ensures backups and recovery work.	Carries out assigned maintenance tasks on a data collection and records every change in the log.	Independently manages data collections, plans maintenance and recovery actions and tests whether a restore actually works.	Determines how data storage is organised for the organisation and decides on platforms, retention methods and continuity requirements.
Data governance and ownership Records who owns and stewards which data, which definitions apply and how decisions about data use are made.	Maintains the register of data owners and definitions as instructed and reports ambiguities to the coordinator.	Assigns ownership for a data domain, has definitions agreed and records decision making about data use traceably.	Establishes the organisation's data governance model and holds executives to their role as owners of data.
GDPR and privacy in practice Knows privacy legislation requirements for own work, which legal basis and retention period apply and recognises when notification is required.	Processes personal data only according to the applicable work instruction and refers doubtful cases to the privacy officer.	Judges in own processes whether a legal basis exists, applies data minimisation and handles requests from data subjects correctly.	Sets privacy policy, decides on high risk processing operations and represents the organisation towards the supervisory authority.
Software testing and quality assurance Devises and runs tests on functionality and edge cases, records findings reproducibly and judges whether a solution can be released.	Works through a test script step by step and records every deviation with a screenshot and the steps taken.	Designs test cases including edge cases, automates repeatable tests and advises on release based on the findings.	Sets the organisation's test strategy and quality requirements and decides when software is not released.

Skill	N1	N3	N5
Selecting and implementing AI applications Selects AI applications based on work need, cost, risk and fit with existing systems and guides adoption through to actual use.	Tests a designated AI application in own work and reports what proved usable and what did not.	Compares applications on requirements, risk and cost, runs a pilot and guides colleagues through adoption.	Determines which AI applications the organisation uses, decides on scaling up and accounts for the choices made.
AI governance and responsible use Applies rules and agreements on permitted AI use, records where AI was used and ensures people remain accountable for decisions.	Uses AI only within permitted use, shares no confidential data and states where AI was used when this is required.	Tests intended AI use against policy and legislation, records the trade off and arranges human oversight of decisions.	Drafts the organisation's AI policy, sets up its oversight and accounts for it to regulators and clients.
Recognising AI risk and bias Recognises where an AI system works systematically skewed or unreliably for groups or situations and names consequences and possible measures.	States that AI output can be skewed by the data used and refers striking outcomes to a colleague.	Examines outcomes per group or situation, identifies where the system falls short and proposes measures or extra checks.	Sets the assessment frameworks for AI fairness and risk and decides whether a system may remain in use.
multidisciplinary collaboration Working with people from other disciplines by learning their language and methods and establishing shared starting points for the joint work.	Asks a colleague from another discipline what technical terms mean and summarises the answer back to them.	Maps the issue with specialists from different disciplines and formulates a joint approach with a clear division of roles.	Sets up lasting forms of collaboration between disciplines and settles conflicts between professional standards on the basis of the organisational interest.
Visual communication and data storytelling Presenting figures and processes in charts, graphs and diagrams that support the message and lead the reader to the correct conclusion.	Creates a chart in an existing template, adds a title and names the unit shown on the axis.	Chooses a fitting chart type per message, removes unnecessary styling and writes the conclusion above the figure.	Sets guidelines for visualisation and dashboard use across the organisation and judges whether displayed figures represent reality fairly.

Skill	N1	N3	N5
Dealing with ambiguity Works on when the assignment, information or division of roles is incomplete or contradictory, makes workable assumptions and tests them once more becomes clear.	Asks what is meant when an assignment is unclear and then works according to the answer given.	Formulates explicit assumptions when information is missing, proceeds on that basis and revises the approach once new data arrives.	Sets direction under high uncertainty, makes the remaining uncertainty explicit and keeps the organisation moving until a decision can be taken.

6 Assessment methods and validity

The validities are the corrected meta-analytic estimates from Sackett, Zhang, Berry and Lievens (2022) in the Journal of Applied Psychology, which replace the older Schmidt and Hunter (1998) figures. The spread matters as much as the mean: with a high SD, validity varies strongly by context.

Assessment method	rho	SD	What it is
Cognitive ability test	0.41	0.14	Standardised test of verbal, numerical or abstract reasoning.
Work sample test	0.33	0.09	The candidate performs a representative work sample under standardised conditions.
Knowledge test	0.40	0.13	Test of declarative job knowledge, usually assembled per client.
On-the-job observation	-	-	Structured workplace observation with a form based on behavioural anchors.
Assessment centre	0.29	0.09	Multiple exercises, multiple assessors, dimension-wise scoring.
Personality questionnaire	0.40	0.15	Self-report of behavioural preferences. At hrnforce the Big Fifty and the 15PF.
Structured interview	0.42	0.19	Behaviour-based interview with fixed questions and a rating scale per question.
Situational Judgement Test	0.12	0.11	Scenario with response options, scored on knowledge or behavioural tendency.

7 Assessment instruments for this job

hrnforce instrument	Skills	What it measures
Work sample test	16	Representative work sample under standardized conditions
Knowledge test (client-specific)	12	Professional knowledge test, customized for each client
Competency Check	6	Behavioral-Level Competency Assessment
Portfolio	5	Evaluation of prior work against fixed criteria
Ability Scan	4	Aptitude Test: Verbal, Numerical, and Abstract Reasoning Skills
Structured interview	4	Behavioral interview with a fixed structure
Big Fifty	2	Personality questionnaire, 150 items, 25 factors

hrmforce instrument	Skills	What it measures
15PF	1	Personality Profile, a shorter alternative to the Big Fifty
Motivation	1	Drivers and sources of motivation
360 Feedback	1	Multiple assessments based on behavioral anchors
Mental Resilience (4C)	1	Mental Resilience According to the 4C Model

8 How to use this profile

1. Agree the profile with the hiring manager before the first candidate is measured. Weights shifted afterwards make the outcome indefensible.
2. Choose the assessment method per skill from the Assessment method column. Combine at most three methods per candidate.
3. Calibrate the assessors in a two-hour session using the anchors in chapter 5. Anchors without rater training deliver little.
4. Fill in the match calculator and discuss the gaps in the development conversation. Type A and type G are selection criteria, not one-quarter development goals.

Source: hrmforce Skills Framework 2.0, version 2.0.0, built 2026-09-08. Crosswalks to ESCO v1.2.1, O*NET 31.0, Lightcast Open Skills, SFIA 9, DigComp 3.0, EntreComp, LifeComp, GreenComp, Allit, WEF Future of Jobs 2025, ISCO-08 and the Dutch CBS occupational classification BRC 2014 edition 2025. This profile is a starting point, not a standard: align it with your own organisation before selecting on it.