Q-Learning Grid World Agent

IEEE-Style Project Documentation and Case Study

Muhammad Haider Abbas | MSc Computer Science & Technology | Ulster University Birmingham Campus



Figure 1. Project visual identity used in the portfolio website.

Abstract

Q-Learning Grid World Agent is documented as a professional portfolio case study by Muhammad Haider Abbas. The project demonstrates practical capability in AI / Machine Learning through a structured problem-solution design, implementation direction, testing plan and future development roadmap. The report explains the business or academic motivation, system requirements, architecture, workflow, quality considerations and expected professional value of the project. It is prepared for use as downloadable portfolio evidence for recruiters, clients and academic reviewers.

Keywords: AI / Machine Learning, Academic Project, ML Developer, RL, Q-Learning, Python, AI

I. Project Profile

Field	Details
Project Title	Q-Learning Grid World Agent
Category	AI / Machine Learning
Year	2026
Role	ML Developer
Status	Academic Project
Tools	Python, NumPy, Matplotlib
Portfolio Slug	q-learning-grid-world-agent

II. Introduction and Problem Context

This document presents the technical and professional documentation for Q-Learning Grid World Agent. The project is positioned as a academic project in the AI / Machine Learning category. The work highlights the author's ability to combine technical implementation, documentation discipline and outcome-focused project thinking. The project summary is: A reinforcement learning project where a Q-learning agent learns optimal movement in a 5×5 grid world with obstacles, sparse rewards and a +5 jump transition.

A. Problem Statement

The agent must learn optimal navigation despite obstacles, sparse rewards and exploration trade-offs.

B. Proposed Solution

Q-learning with epsilon-greedy exploration is used to compare learning behaviour across different exploration values.

III. Aim and Objectives

To implement and evaluate a Q-learning agent in a deterministic 5x5 grid world with obstacles, sparse rewards and a special jump transition.

- Define the state space, action space, reward model and terminal state.
- Implement Q-table learning with epsilon-greedy exploration.
- Compare exploration values and learning rates.
- Visualise reward progression, steps per episode and learned policy.

IV. Requirements Analysis

ID	Requirement	Type	Priority
FR-01	5x5 grid environment	Functional / project requirement	Required
FR-02	Obstacle handling	Functional / project requirement	Required
FR-03	Terminal reward +10	Functional / project requirement	Required
FR-04	Jump reward +5	Functional / project requirement	Required
FR-05	Q-learning update	Functional / project requirement	Required
FR-06	Epsilon comparison charts	Functional / project requirement	Required
NFR-01	Usability and clear portfolio presentation	Non-functional	Required
NFR-02	Maintainability and future extensibility	Non-functional	Required
NFR-03	Professional documentation and evidence quality	Non-functional	Required

V. System Architecture

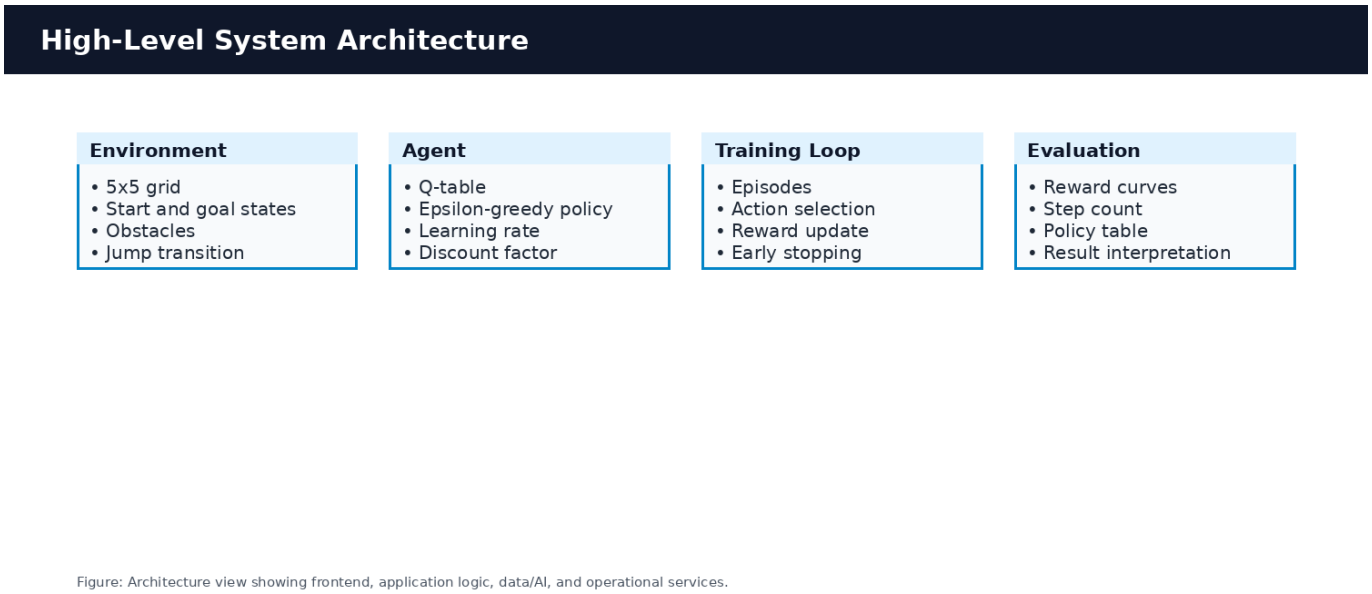


Figure: Architecture view showing frontend, application logic, data/AI, and operational services.

Figure 2. High-level architecture or production pipeline for the project.

The architecture separates user-facing presentation, project logic, data or model processing, and operational output. This structure makes the project easier to explain, test and extend inside a professional portfolio environment.

VI. Workflow Design

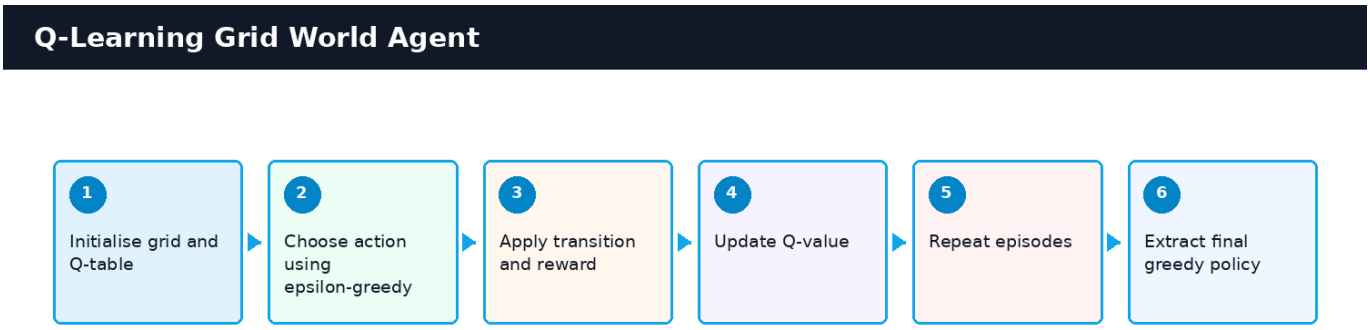


Figure: Generated project workflow / architecture artefact for portfolio documentation.

Figure 3. Project workflow from input or user action to final output.

VII. Data, Components and Implementation View

Component	Layer	Responsibilities	Notes
C-01	Environment	5x5 grid, Start and goal states, Obstacles, Jump transition	Core component group
C-02	Agent	Q-table, Epsilon-greedy policy, Learning rate, Discount factor	Core component group
C-03	Training Loop	Episodes, Action selection, Reward update, Early stopping	Core component group
C-04	Evaluation	Reward curves, Step count, Policy table, Result interpretation	Core component group

Implementation should be documented through screenshots, code snippets, database schema extracts, test outputs or demonstration media when available. For deployed or live systems, screenshots and logs should be attached as appendices to increase evidence strength.

VIII. Testing and Validation Strategy

Test Area	Test Method	Expected Result
Transition rules	Move into wall/obstacle	Agent remains valid
Reward logic	Reach jump/terminal	Correct reward returned
Q-update	Run controlled episode	Q-table changes
Evaluation	Plot rewards and steps	Results match experiment design

The testing strategy focuses on validating the main user or system journey, confirming that core requirements behave as expected, and recording limitations clearly. This supports honest portfolio evidence and avoids overclaiming features that are still prototypes.

IX. Security, Ethics and Professional Considerations

- Avoid presenting AI recommendations as guaranteed truth.
- Protect user data and avoid uploading sensitive CV, health or behavioural data without consent.
- Use evaluation metrics and explain limitations clearly.

X. Results, Outcome and Future Work

A. Outcome

A clear academic AI project showing practical reinforcement learning understanding.

B. Future Enhancements

Add visual simulation, animated training steps and comparison charts.

Roadmap Stage	Enhancement Direction
Short term	Add missing screenshots, demo links and downloadable supporting files.
Medium term	Connect live/source URLs and replace placeholder evidence with real artefacts.
Long term	Use analytics, user feedback or formal evaluation results to strengthen the project case study.

XI. References

[1] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed., MIT Press, 2018.

[2] C. J. C. H. Watkins and P. Dayan, "Q-learning," Machine Learning, vol. 8, no. 3-4, pp. 279-292, 1992.

[3] Muhammad Haider Abbas, "Deep Learning and Its Application Assignment," uploaded coursework artefact, 2026.

Appendix A. Portfolio Download Mapping

Item	Value
Recommended PDF path	docs/q-learning-grid-world-agent-project-documentation.pdf
Recommended Word path	docs/q-learning-grid-world-agent-project-documentation.docx
Admin field: Documentation	docs/q-learning-grid-world-agent-project-documentation.pdf
Admin field: Extra Project File	docs/q-learning-grid-world-agent-project-documentation.docx