|  |
| --- |
|  |

|  |
| --- |
| **CYBER CREW.TECHNICAL DOCUMENTATION** |
| AI KNOWLEDGE CHATBOT |
|  |
| NHLALUKO MALULEKE; KARABO MOKWANA; PRISCILIA MACAMO; NKOSINATHI MLALAZI; NOSIBUSISO MHLOHLO. |
|  |
|  |
| **5/19/2025** |
|  |

TABLE OF CONTENTS

Objective 2

Core Concepts 2 - 3

Types of learning in Machine Learning 3

Real-World Applications 4

Ethical Considerations in AI 4

Summary 5 - 7

AI C**hat**bot Knowledge Base

# Objective

The objective is to develop a publicly accessible AI chatbot that educates users on AI concepts through natural, context-aware conversations, while also recommending relevant learning resources and citing specific bootcamp course materials for deeper understanding.

# Core Concepts

## Artificial Intelligence (AI)

AI is the ability of computers or machines to mimic human intelligence, including learning, reasoning, and problem-solving.

## Machine Learning (ML)

Machine Learning is a subset of AI that involves training algorithms to make predictions or decisions based on data.

## Deep Learning

Deep Learning is a type of Machine Learning that leverages neural networks with many layers ('deep' networks) to process complex patterns in large datasets.

## Natural Language Processing (NLP)

NLP is a field of AI that gives machines the ability to read, understand, and derive meaning from human languages.

## Neural Networks

Neural Networks are a series of algorithms that attempt to recognize underlining relationships in asset of data through a process that mimics the way the human brain operates.

## Computer Vision

Computer Vision a field of AI that enables computers to interpret and make decisions based on visual data.

## Large Language Models (LLMs)

LLMs are advanced AI models trained on vast amounts pot text data to perform functions like text generation and question-answering

## Training Data

Training data refers to the labelled or unlabelled dataset used to build and train machine learning or AI models. The model uses this data to adjust its internal parameters to make accurate predictions or decisions.

Reinforcement learning is where an AI agent learns by interacting with its environment and receiving rewards or penalties.

## Bias

Bias in AI refers to systematic and unfair discrimination that occurs when an artificial intelligence system produces prejudiced results due to flawed data, algorithms, or underlying assumptions. This bias can manifest in many ways, such as favouring one demographic group over another, reinforcing stereotypes, or misrepresenting information. AI systems learn from data provided by humans, and if that data reflects historical inequalities, cultural stereotypes, or imbalanced representation, the AI may learn and replicate those patterns—sometimes at scale.

# Types of Learning in Machine Learning

## Supervised Learning

Uses labeled data to train the model. The model learns to map input to known output.

## Unsupervised Learning

Involves training a model on data without labels, helping it find hidden patterns or groupings.

Supervised learning is a type of ML where the model learns from labelled training data to predict outcomes for new data.

## Reinforcement Learning

Involves an agent learning to interact with its environment through rewards or penalties.

# 

# Real-World Applications

AI is used in industries like healthcare, finance, transport, and more.

## Healthcare

Disease detection via imaging, personalized treatment planning, predictive analytics.

## Finance

Fraud detection, algorithmic trading, credit scoring, customer service automation.

## Retail

Personalized recommendations, inventory optimization, chatbots, demand forecasting.

## Education

Adaptive learning, virtual tutors, automated grading and admin support.

# Ethical Considerations in AI

## Bias in AI

AI systems may inherit human biases from training data, causing unfair outcomes.

## Privacy Concerns

AI systems may misuse or inadequately protect user data.

## Transparency and Explainability

Understanding AI decision-making is crucial for trust and fairness.

## Job Displacement

Automation can replace human jobs, raising concerns over employment.

# Summary

This documentation provides a foundational understanding of key AI concepts. It serves as a reference for building and understanding AI c**hat**bot technologies.

**What is the objective of this AI chatbot project?**

To develop an AI chatbot that can:

* Explain core AI concepts.
* Sustain contextual follow-ups (at least 3 deep).
* Recommend related learning materials (e.g., course modules).
* Cite relevant bootcamp materials.
* Be accessible via a public URL.

### ****How will the chatbot handle follow-up questions?****

### The chatbot will:

* Use a **conversation memory buffer** to retain the context.
* Track on-going topics using **semantic similarity**.

### ****How will the chatbot provide further learning recommendations?**** It will:

* Reference a **lookup table** of course modules mapped to specific AI concepts.
* Offer module suggestions post-response.

### ****Will the chatbot provide citations to bootcamp materials?****

* Module names
* Slide numbers
* Lecture recordings

### ****How does the chatbot support cross-linking between concepts?****

### It will use a **graph-based structure** to detect and suggest related terms. eg:

### Neural Networks are closely related to Deep Learning and LLMs.

### ****What are the key deliverables for this project?****

1. **Deployed Chatbot**
   * Hosted publicly (e.g., on GitHub).
2. **Technical Documentation**
   * Covers system architecture, user flows, deployment, and evaluation.

### ****What are the evaluation criteria for the chatbot?****

| **Criteria** | **Description** |
| --- | --- |
| Accuracy | Explanations match bootcamp curricula and are factually correct. |
| Conversational Flow | Natural, logical, and human-like conversations. |
| Edge Case Handling | Robust responses to vague, ambiguous, or tricky questions. |
| Feedback Mechanism | Let users rate or respond to chatbot replies for future improvements. |
| Documentation Quality | Clear, professional-level tech documentation for developers and stakeholders. |

### ****What does the system architecture look like?****

User ↔ Frontend Interface (Chat UI)

↕

Chat Engine (LLM-based API)

↕

Memory Handler (Session + Context Management)

↕

Knowledge Base (AI Concepts + Course Modules + Citations)

↕

Recommendation Engine (Further Learning + Cross-links)

### ****What optional enhancements could improve the chatbot?****

* **Quiz Generation**: Generate quizzes from prior conversations.
* **Multimodal Support**: Accept voice queries or display visual aids (like course slides).
* **Admin Panel**: For updating course metadata and response mappings.