Artificial Intelligence/Machine Learning Policy for SMEs

Prepared by Dr Howard Haughton CMath FIMA
17 August 2023

1. Introduction

1.1 Purpose

The purpose of this policy is to establish guidelines and standards for the implementation of artificial intelligence/machine learning (AI/ML) technologies within our organisation to ensure data privacy, ethical behaviour, and the responsible use of AI/ML in solving problems and delivering value. The policy is based on certification standards developed by the Alliance for Data Science Professionals which can be accessed here https://alliancefordatascienceprofessionals.com/.

1.1 Scope

This policy applies to all employees, contractors, and stakeholders involved in the development, deployment, and utilisation of AI/ML solutions within our organisation.

2. Problem Definition and Communication with Stakeholders

This section centres around effective stakeholder engagement and problem definition to drive successful outcomes. It involves identifying project requirements, articulating clear problem statements, and evaluating assumptions and biases. It highlights risk assessment, sector/domain knowledge, and showcasing the value of data science. Strong relationship management, clear communication with diverse audiences, and considering human factors in implementing data-driven solutions are emphasised. A key consideration is whether AI/ML or some other approach would be more appropriate.

2.1 Problem Definition

Having an accurate understanding of the problem definition is central to the success of any AI/ML model. Effective problem definition means that we can align project goals with organisational and stakeholder objectives.

2.1.1 Project Requirements Elicitation

Project requirements elicitation is as much an art as it is science. It involves understanding human factors such as emotions, motivation, culture, and communication and logical and domain specific analysis of end-users documented requirements, assuming they exist. Techniques employed from traditional software development for requirements engineering provide a systematic means of capturing end-user expectations for model development and should be preferred over informal non-documented processes. Once documented, these requirements should be version controlled and appropriate change control methodologies put in place to agree any updates with all stakeholders.

A key objective of requirements elicitation is not just to document user's functional requirements but to also understand the constraints under which such requirements must be achieved (non-functional requirements e.g., performance, accuracy, reliability). Collectively, these form the success criteria for an AI/ML application. It provides a basis for developers and end-users to have a common understanding of the problem(s) to be solved as well as the potential issues that could arise from the specification due to ambiguities or other factors.

2.1.2 Risk Assessment and Domain Knowledge

As for traditional software development, model developers and end-users should assess the risks associated with the problem domain and agree on proposed solutions. Model development teams should be able to demonstrate sector/domain knowledge and understand how data science can deliver value in specific sectors/domains. Consideration of the likelihood and severity of risks arising from the use of AI/ML should inform the level of rigour applied at the various stages of model development. It is encouraged, at an early stage, that this analysis also takes into account whether the use of AI/ML would be the most suitable means of addressing the business need required of a software solution as alternative strategies might be more appropriate.

2.2 Project Management Approaches

In our organisation, it is understood that effective project management is vital in AI/ML development, ensuring efficient resource allocation, risk management, collaboration, and adherence to timelines. It defines clear scopes, manages stakeholder expectations, and facilitates continuous improvement, maximising benefits while minimising challenges.

2.2.1 Solution Identification

The choice of solution can have an impact on the effectiveness and usability of an application within a business domain. Hence, developers should identify viable solutions based on project requirements, available data and taking into consideration external factors such as laws/regulations and the need for transparency and explainability of a model's results. Consideration should be given to the risks associated with whether data is to be labelled or unlabelled and use the approach best aligned with the objectives/goals of the model and available resources after accounting for the risks. Once a decision is made, developers should discuss the implications to technical and non-technical stakeholders and obtain agreement on the most appropriate solution.

2.2.2 Technical and Project Management Methodologies

Utilising a formalised structured framework to plan, monitor and control project activities (including the use of version and configuration management tools) provides our organisation with a basis for successfully delivering an AI/ML project. This framework should define clear project objectives, scope, timelines, and milestones as well as required resources, risks and controls and provide for ongoing monitoring of progress. This promotes efficient collaboration between development team and other stakeholders.

2.3 Relationship Management

Relationship management is crucial for our organisation as it enables effective collaboration, communication, and understanding with our stakeholders. Building and maintaining strong relationships with all stakeholders is essential for the success of our data science projects.

2.3.1 Effective Communication

Model developers should be able to communicate effectively with diverse audiences, including technical colleagues, subject matter experts, and leadership when discussing any aspect of the AI/ML lifecycle. Whilst not being prescriptive, evidence of effectiveness could include formal agreement on artefacts such as requirements specification, applicable test cases, constraints/restrictions on model use etc. Training and user guides should be provided to show how the new solution will be deployed and/or embedded into the company.

2.3.2 Stakeholder Expectation Management

As for traditional software development, an AI/ML application is likely to have several potentially diverse stakeholders. These stakeholders will not, necessarily, always have the same opinions concerning a particular application. It is expected that model developers will be able to effectively manage the expectations of diverse stakeholders with conflicting priorities and mediate equitable solutions to ensure stakeholder satisfaction.

3. Data Privacy and Stewardship

This section discusses the importance of data privacy and stewardship in ensuring the security and protection of data throughout its lifecycle to comply with legal and regulatory requirements.

3.1 Ensuring the protection of personal and sensitive data

Ensuring the protection of personal and sensitive data is crucial for our organisation as it builds trust with customers and stakeholders, enhances reputation, and mitigates the risks of data breaches, unauthorised access, financial losses, legal consequences, and reputational damage.

3.1.1 Risk Assessment and Data Protection Policies and Procedures

To ensure the protection of personal and sensitive data, at a minimum, the following three activities should be undertaken:

1. Regular risk assessments to identify potential vulnerabilities and risks related to personal and sensitive data. This can be undertaken either as part of dedicated risk assessments related to data protection or as part of wider organisational risk management initiatives e.g., enterprise risk assessment.
2. Formalisation and implementation of data protection policies and procedures that align with legal and regulatory requirements e.g., The UK Data Protection Act 2018 and General Data Protection Regulation or standards such as ISO/IEC 27001.
3. Employees are trained on data protection practices and adhere to organisational policies.

This is covered in our company by separate policies <LIST POLICIES HERE IF APPLICABLE OTHERWISE DELETE>.

3.1.2 Safe and Secure Management of Sensitive Data, Models, and Infrastructures

Once risk assessment and data policies have been implemented AI/ML model developers/users should ensure that data is securely stored and implement role-based access controls to protect sensitive data and models from unauthorised access or disclosure. If security vulnerabilities are revealed, then these should be immediately addressed and a process involving regular updates to system security and infrastructure should be initiated. In any event, access to all sensitive data and models should be monitored and logged to detect and respond to any unauthorised activities.

3.1.3 Application of Data Controls

As far as practically possible, data controls, such as encryption, (pseudo) anonymisation, and synthetic data generation, to protect privacy and mitigate risks, should be employed. Secure methods for data sharing and distribution should be used to ensure that data is only accessible to authorised individuals or entities. Mechanisms should be implemented for data anonymisation and de-identification when necessary.

3.1.4 Risk Management around Environment and Infrastructure

AI/ML applications are software systems often embedded as part of other decision support systems. In managing the operational risks of these applications, it is also necessary to consider risks related to the IT infrastructure and environment. Specifically, and prior to an application's use in supporting business decisions, risk management professionals should ensure that:

1. Risk management processes have been established to identify and address potential risks related to the environment and infrastructure used for AI/ML activities. The responsibility for implementing such risk management activities should be consistent with organisational risk management policies.
2. The risk management process incorporates identification, analysis, evaluation, and treatment of risks associated with data storage, network security, and infrastructure vulnerabilities.
3. Procedures exist for data and systems backups and that disaster recovery plans exist to ensure data availability and integrity in the eventuality of a failure in any component supporting the use of the AI/ML application.

3.2 Managing Sensitive Data

Effective management of sensitive data is essential for our organisation. It safeguards confidential information, ensures compliance, and builds trust with customers. Mishandling data can result in reputational damage, legal issues, and financial losses. By implementing robust data protection measures, we demonstrate our commitment to privacy and security, enhancing our overall success and resilience in a data-driven business environment.

3.2.1 Integrity and Legal Compliance

Employees are expected to act in a responsible and professional manner regarding conformity to legal and regulatory requirements regarding the handling of sensitive data. They are expected to be familiar with relevant legal and regulatory procedures for responding to potential data loss or breaches.

3.3 Data Stewardship and Standards

Data stewardship and standards are vital for ensuring the reliability, consistency, and interoperability of data within our organisation. By establishing data stewardship practices and adhering to industry standards, we promote data quality, integrity, and trustworthiness. This enables effective data sharing, collaboration, and decision-making across different teams and systems. Furthermore, data stewardship and standards support compliance with regulatory requirements and facilitate seamless integration of data from diverse sources. Ultimately, they contribute to the organisation's ability to make informed decisions, drive innovation, and derive meaningful insights from data.

This is covered in our company by our Data and Knowledge Management Policies <INSERT REFERENCES HERE IF APPLICABLE OTHERWISE DELETE>.

3.3.1 Incorporating FAIR Guiding Principles

As far as is practically possible and where it is legally and ethically appropriate the organisation should incorporate the FAIR Guiding Principles for scientific data management and stewardship into our data practices. These principles can be accessed here https://www.go-fair.org/fair-principles/. From our organisational perspective, the FAIR guiding principles could aid us in enhancing the findability, accessibility, interoperability, and reliability of our data sets. This provides additional value to our stakeholders as well as enabling us to further leverage our potential to drive innovation, collaboration and more-informed decision making.

3.3.2 Technical Standards, Regulation, and Governance

Technical standards, regulation, and governance are essential for our organisation to ensure effective data management. Adhering to established standards promotes data consistency and interoperability, while compliance with regulations protects privacy rights and mitigates penalties. Robust governance frameworks provide oversight and accountability. By prioritising these aspects, we create a secure and trustworthy data environment, enhance data sharing capabilities, and build stakeholder confidence.

4. Definition, Acquisition, Engineering, Architecture, Storage, and Curation

This section covers the entire data lifecycle, including definition, acquisition, engineering, architecture, storage, and curation. It focuses on sourcing, analysing, and constructing datasets, ensuring data quality, and addressing infrastructure requirements. It highlights the importance of secure deployment and scalable delivery of data products, considering data provenance, feature modelling, and different data types.

4.1 Data Collection and Management

Effective data collection and management are vital for our organisation as they enable informed decision-making, drive innovation, and enhance operational performance. By collecting and managing data efficiently, we gain valuable insights into customer behaviour, market trends, logistics, and process optimisation. Moreover, it ensures compliance with regulations, maintains data accuracy, integrity, security, and privacy. Through these practices, we maximise the potential of our data assets, stay ahead in the market, and improve the likelihood of achieving sustainable growth.

4.1.1 Source and Access Data

When sourcing data, identify and access appropriate data sources that align with the problem at hand and critically analyse the availability, amount and relevance of data and resources to meet project requirements. If data is being sourced externally you must consider whether you have the appropriate copyright and license conditions for your use case.

4.1.2 Data Evaluation and Synthesis

Effective data evaluation and synthesis are essential for our organisation as they enable us to gain valuable insights, make informed decisions, and drive positive outcomes. By analysing and synthesising data, we can uncover patterns, trends, and correlations that help us optimise processes, develop strategies, and identify growth opportunities. It also allows us to measure performance and track progress towards our goals, ensuring continuous improvement and competitiveness in the dynamic business environment. To achieve this AI/ML implementors should:

1. Evaluate and synthesise data from multiple sources to ensure comprehensive analysis.
2. Follow data provenance processes to maintain data integrity and traceability.

4.1.3 Data Characteristics and Infrastructure Requirements

As our business environment changes, our data and infrastructure capacity might also need to maintain pace with these dynamics. Key to this, is understanding the characteristics of data including:

- Volume: referring to the amount of data generated or stored for some aspect of AI/ML
- Velocity: referring to the speed at which data is generated, processed, or transferred
- Variety: referring to the diversity of data formats and types e.g., structured, semi-structured or unstructured

Understanding the characteristics of data helps our organisation to apply data management techniques tailored to the specific data types and formats ensuring effective and efficient storage, retrieval, and usage.

4.2 Data Engineering

Data engineering involves creating and maintaining the systems and processes to collect, transform, and store data for effective analysis and decision-making. It plays a vital role in our organisation, enabling us to leverage data for informed decision-making and strategic progress. Efficient data management and processing unlock valuable insights, improve operational efficiency (for example, using the knowledge of the existence of data drifts or changes to the data population as an indicator of the possible need to re-calibrate an AI/ML model), and drive innovation. We prioritise data availability, reliability, and integrity through activities such as data integration, transformation, and pipeline development, supported by robust data architectures and infrastructure.

4.2.1 Constructing Data Sets

Prior to building any AI/ML model it is imperative that model developers gain a comprehensive understanding of the data landscape. Specifically, constructing data sets by linking and integrating data from multiple disparate sources and performing data profiling enables us to uncover valuable relationships, patterns, and insights which might not be apparent when examining individual data sources. Such analysis provides insight into the completeness, accuracy, and consistency of data. Specifically, integrating data from diverse sources helps to provide a more comprehensive view of a business domain and aids us in reducing the likelihood of bias. Of particular importance, is an assurance that the analysis conducted serves to highlight where bias might exist and how it is addressed prior to using data in a model. Data profiling and data exploration enables us to better assess the quality and reliability of the data including error and inconsistency detection or gaps that could affect the accuracy and validity of our models.

A key outcome of data analysis is an assessment of whether the data is sufficient and appropriate to support the development of an AI/ML model to address the business requirements. This assessment should be undertaken prior to model building.

4.2.2 Handling Missing Data

Depending on the context, it might be necessary to address the issue of missing data. In such cases developers should implement principled inclusion/exclusion criteria and imputation methods to handle missing data effectively. The process adopted should be documented and, where possible, allow for reproducibility of the results. A key objective in any approach adopted should be to minimise the introduction of bias and/or the potential for infringement of any law/regulation.

4.2.3 Data Curation and Quality Controls

Data curation and quality control form an important part of the overall quality assurance of AI/ML models within our organisation. Our approach to curation is systemised in that it documents the data lifecycle including creating/sourcing, organising, and maintaining. This should be undertaken in a manner that allows individuals to assess the provenance of data as well as to verify its integrity and accuracy.

4.2.4 Infrastructure Selection

As for any regular software development project, prior to the start of development, our organisation should identify the most appropriate infrastructure solutions, considering factors such as scalability, security, and cost-effectiveness. Depending on the business/project needs and considering our current and prospective technology capacity it might also be necessary to evaluate and select between make/buy and cloud and on-premises options.

5. Problem Solving, Analysis, Statistical Modelling, Visualisation

This section focuses on problem-solving, analysis, statistical modelling, and visualisation in data science. It involves applying technical solutions based on requirements and available data. Emphasis is on data preparation, feature modelling, evaluation metrics, and systematic analysis. It highlights the use of suitable tools and effective visualisation for communicating findings.

5.1 Identifying Technical Solutions Approaches

Identifying technical solutions and appropriate approaches is important because it allows us to effectively address the challenges and complexities of projects requiring the use of data science within our organisation.

5.1.1 Solution Identification

The choice of model to use can have an impact on the effectiveness and usability of an application within a business domain. Hence, developers should identify viable solutions based on project requirements, available data and taking into consideration external factors such as laws/regulations and the need for transparency and explainability of a model's results. Once a decision is made, developers should discuss the implications to technical and non-technical stakeholders and obtain agreement on the most appropriate solution.

5.1.2 Technical and Project Management Methodologies

Utilising a formalised structured framework to plan, monitor and control project activities provides our organisation with a basis for successfully delivering an AI/ML project. This framework should define clear project objectives, scope, timelines, and milestones as well as required resources, risks and controls and provide for ongoing monitoring of progress. This promotes efficient collaboration between development team and other stakeholders.

5.2 Data Preparation and Feature Modelling

The basic building blocks of any AI/ML model are the features used to make predictions. In turn, these features are based on curated data.

5.2.1 Solution Selection and Evaluation

Developers should employ appropriate statistical and machine learning approaches for data analysis. They should define and evaluate suitable evaluation metrics, including computational performance and accuracy. For purposes of reproducibility, details of approaches used and the datasets that are applied should be documented.

For transparency and model explainability evaluation may need to go further by understanding model sensitivity to input features, any unwanted biases to slices of the data and the range of answer variation as the inputs are varied. Customers may require model explanations and results in visualisable manner for effective communication. Of particular importance, is assessing the implications of bias in both the data inputs and outputs of an AI/ML model.

5.2.2 Data Manipulation and Feature Creation

Data comes in several forms e.g., numerical, categorical, or textual and understanding its characteristics allows us to apply the most appropriate methods of analysis such as normalisation, scaling or encoding. Effective manipulating and transformation can help to extract meaningful insights, identify patterns, and uncover relationships within data. This process enhances the predictive power of models.

5.3 Data Analysis, Visualisation and Model Building

Data analysis, visualisation, and model building are vital for deriving insights, making informed decisions, and improving business outcomes through statistical and machine learning techniques.

5.3.1 Applying Statistical and Machine Learning Approaches

Model developers should employ appropriate statistical and machine learning techniques to solve the problem at hand. Measures of appropriateness go beyond just focussing on accuracy but should also account for ethical, legal/regulatory, and other considerations. It should also be the case that models are developed in a programming language for which there is a high degree of support and usage within the world of software development.

Developers should adopt a systematic approach to exploratory data analysis, managing ambiguity and uncertainty and utilise suitable visualisation tools to communicate and demonstrate results. Part of this systematic approach is ensuring that any evaluation and visualisation scripts that have been used to choose between models are also considered part of the solution software pipeline, and therefore are version controlled and recorded.

5.4 Software Development practices

Methods employed in the development of AI/ML systems should be based on best practices covering all phases of the model development life cycle.

5.5 Deployment

This is a critical phase in the development of an AI/ML model as it implies that the model is ready for operational use by end-users/stakeholders. Deployment typically involves the implementation, integration and delivery of solutions that can aid stakeholders to solve specific business problems.

5.5.1 Planning Deployment and End-User Engagement

The deployment of all AI/ML models should be undertaken in collaboration with end-users as for any traditional software development project. Prior to deployment, it is expected that developers would have undertaken rigorous testing of applications employing relevant metrics of accuracy and goodness of fit along, where necessary, with unit, system, integration, and other pre-user testing. All such testing should be documented including tests applied and results.

Only once testing by developers are complete should an application be given to end-users for testing. The purpose of end-user testing is to ensure the application is fit for use in real-world scenarios. Feedback from users might necessitate refinements of applications to address any issues or limitations identified. The results of user acceptance tests should also be documented and retained.

5.5.2 Monitoring, Maintenance and End of Life

Once an application has been developed, developers (and or other relevant staff) should ensure that formalised processes for monitoring and maintenance are in place to ensure the safe (e.g., no harmful effects of bias) secure, stable, and scalable operation of AI/ML data products.

Consideration must be given to how any solution will be retired from usage at end-of-life or what metrics would signal the need to replace the system in the future.

5.5.3 Accountability, Registering models and Approval

An established process is necessary to pinpoint the individuals accountable for decisions or complications arising from the software, which is particularly intricate in the context of ML/AI. This protocol should be precisely delineated. During the deployment of software and models, pivotal particulars like the model's specifics, dataset particulars, the personnel held accountable and responsible, along with the approvers, must be comprehensively documented and registered. Information detailed (among others) would include dataset sources, size, architecture, performance metrics, intended user base, usage limitations, ethical considerations, updates, contact information, and external dependencies. By incorporating these points, transparency, accountability, and informed decision-making regarding AI/ML models are significantly enhanced.

6. Evaluation and Reflection

The focus of this section centres on ongoing improvement, ethical awareness, sustainability, and continuous professional development. It entails monitoring project performance and outcomes, deriving insights from experiences, and engaging in collaborative project evaluations. Crucial elements encompass maintaining data integrity, mitigating biases, adhering to legal and regulatory obligations with integrity, and upholding ethical principles to ensure secure data and AI usage. Moreover, this section emphasises the significance of open science, reproducible research, upholding high technical standards, and adhering to best practices in software development.

6.1 Project Evaluation

Project evaluation holds significant importance for our organisation as it enables us to assess the success and impact of our initiatives. By conducting comprehensive evaluations, we can identify areas of improvement, optimise resource allocation, and enhance our project outcomes.

6.1.1 Ongoing Performance Monitoring and Improvement

For every project, there should be continuous monitoring of project performance and outcomes, ensuring that they align with the defined objectives. Every opportunity should be used to identify and apply lessons learned to future projects. Such retrospectives should be undertaken individually and in conjunction with all relevant stakeholders.

6.2 Ethical Behaviour

In developing and using AI/ML models, our organisation expects individuals to conduct themselves in a manner that aligns with high ethical standards.

6.2.1 Data Risks and Integrity

Developers/users are expected to identify and manage risks related to erroneous and biased data and act with integrity, adhering to legal and regulatory requirements regarding data privacy and usage.

6.2.2 Ethical Use of Data and AI Technologies

Developers/users are expected to uphold principles of ethical and safe use of data and AI technologies and implement data use procedures to ensure that sensitive data is only used for its agreed-upon purpose.

6.2.3 Data Retention Strategies

All employees involved in the use of data for AI/ML are required to conform to data retention strategies in line with regulatory and legal requirements and ensure that data is retained only for as long as necessary and securely disposed of when no longer needed.

6.3 Sustainability and Best Practices

Persons involved in the development of AI/ML models should, where feasible and appropriate, incorporate the principles of open science and reproducible research into the organisation's practices.

6.4 Reflective Practice and Ongoing Development

Reflective practice and ongoing development are vital for continuous improvement in our organisation and we encourage all of our staff to undertake Continued Professional Development. They enable critical evaluation, learning from experiences, and staying updated with advancements. It fosters innovation, collaboration, and maintains a competitive edge for our organisation.

6.4.1 Learning from Experience, Knowledge sharing and keeping up to date

Model developers should engage in self-assessment of responses to work-based situations to learn from experience in the development/application of AI/ML systems. This experience, where feasible, should be shared with others in the organisation and beyond, if deemed necessary and subject to organisational protocols. Developers should also identify learning opportunities and actively pursue ongoing professional development in data science and to keep abreast of latest developments in industry and academia.