

### PP Coursework Cover Sheet

Please use the table below as your cover sheet for the 1<sup>st</sup> page of the submission. The sheet should be before the cover/title page of your submission.

Programme	
Module name	
Schedule Term	
Student Reference Number (SRN)	
Report/Assignment Title	
Date of Submission <i>(Please attach the confirmation of any extension received)</i>	

#### Declaration of Original Work:

- I hereby declare that I have read and understood BPP's regulations on plagiarism and that this is my original work, and that I have researched, undertaken, completed and submitted in accordance with the requirements of BPP School of Technology.
  - I declare that where I have used any AI tools, it was for the following reasons (highlight as appropriate):
    - To create an assignment plan
    - To create a draft
    - To correct language errors
    - Other (please describe)
- 
- I have copied & retained for BPP University's reference, all AI prompts used in the creation of AI content and all AI-generated responses in support of my assignment and attached relevant evidence in the appendices.
  - I understand that I may be required to participate in a *viva voce*, where I will be questioned on any aspect of my assignment, including key concepts, theories, examples used, & any sources included.
  - The word count, excluding contents table, bibliography and appendices, is \_\_\_\_ words.

**Student Reference Number:**

**Date:**

By submitting this coursework you agree to all rules and regulations of BPP regarding assessments and awards for programmes. **Please note, submission is your declaration you are fit to sit.**

BPP University reserves the right to use all submitted work for educational purposes and may request that work be published for a wider audience.

**BPP School of Technology**

# DTSS Level 7 Data Analytics integrated degree apprenticeship (MSc Applied Data Analytics)

## Machine Learning and A.I.

### Coursework Assessment Brief

Submission mode: **Turnitin online access**

## 1. Assessment Brief

This assessment brief gives you an overview of the formative and summative assessments that are part of this module. The learning outcomes below will be tested in the assessment contained in this brief.

### 1.1. Module Learning Outcomes (LOs)

**LO1:** Critically evaluate machine learning algorithms and their applicability to business problems

**LO2:** Formulate an approach to implementing a machine learning solution

**LO3:** Critically evaluate the success of a machine learning models in solving business problems

### 1.2. Assessment Overview

Two-part assessment which is split **50% for Part 1** and **50% for Part 2**.

- **Formative Submission deadline for Parts 1 and 2:** Assessment to be submitted by **the time shown on the module page on the Hub** in to the respective Turnitin Dropboxes available on the Module page in the Hub.
- **Summative Submission deadline for Parts 1 and 2:** Assessment to be submitted by **the time shown on the module page on the Hub** in to the respective Turnitin Dropboxes available on the Module page in the Hub.

**Please Note:** ensure you read the general assessment guidance at the end of this document.

## Assessment Structure/Guidance

### Part 1: Machine Learning Models (50%)

This first part of the assessment specifically covers:

**LO2:** Formulate an approach to implementing a machine learning solution

**LO3:** Critically evaluate the success of a machine learning models in solving business problems

This assessment is designed to evaluate your ability to critically analyse and implement machine learning models for image classification using the CIFAR-10 dataset. You will explore, implement, and compare the performance of different models, showcasing your understanding of their strengths, limitations, and application.

**Data Description:** You are required to use the CIFAR-10 dataset to perform image classification. The dataset consists of 60,000 32x32 colour images in 10 classes, with 6,000 images per class.

**Submission format:** a well-documented Interactive Python Notebook (.ipynb) containing:

- Code implementation.
- Visualisations and tables summarising results.
- Clear and concise explanations of each step and findings.

A basic template is provided to help you get started, and can be accessed [here](#) or downloaded from the Hub. The use of the template is not compulsory, however ensure that in your submission you clearly indicate your SRN and include all the required content.

#### Task overview

You will develop an interactive Python notebook including the following:

(A) Exploratory Data Analysis and Preprocessing: Carry out a concise EDA to familiarise yourself with the CIFAR-10 dataset. Some questions to answer are:

##### 1. Dataset Overview

- What is the shape of the training and testing datasets?
- What are the data types of the images and labels?
- How many classes are present in the dataset? What are their names?

##### 2. Distribution of Classes

- What is the distribution of samples across the classes? Are the classes balanced or imbalanced?

##### 3. Image Properties

- What is the size (dimensions) of each image?

- What is the range of pixel values in the images? Are they normalized (e.g., 0-1 or 0-255)?

#### 4. Visualization

- Can we visualize a few random images from each class to understand their features?
- What do these images tell us about the variation within and between classes?

(B) Test at least one clustering algorithm such as K-means. Reduce the dimensionality of the dataset using PCA before performing the clustering.

(C) Split the dataset into training and testing set. Implement a classification model based on each one of the following techniques

- a. Decision tree
- b. Random forest
- c. SVM
- d. Research and implement another method such as (but not limited to) k-Nearest Neighbours or XGBoost.

Provide detailed explanations, evaluations, and comments on the outputs of each model, demonstrating a comprehensive understanding of machine learning concepts.

(D) Neural Networks:

- a. implement and compare the performance of a shallow network and a suitable CNN
- b. Research and implement a pre-trained model such as (but not limited to) ImageNet

(E) Hyperparameter Optimization: Select one model and demonstrate the possibility of further improving its performance using hyperparameter optimization techniques.

(F) Results Summary and Discussion: Provide a written and visual summary of the results obtained from the model evaluations. Clearly present the performance metrics and any insights gained from the analysis.

(G) References: the key decisions in previous sections must be supported by evidence from the literature. Include primary research and relevant case studies.

### **Part 1 Formative Submission**

Submission for Part 1 of this assessment will be made in the form of an .ipynb (interactive Python notebook) covering tasks A, B, C, F and G with the relevant code and mark up text (800 words equivalent total). The contents of part F and G will refer only to the completed task.

**Ensure you train and evaluate your model before submitting as the examiner will not train every model received.**

### **Part 1 Summative Submission**

Submission for Part 1 of this assessment will be made in the form of an .ipynb (interactive Python notebook) with the relevant code and mark up text (1800 words equivalent total). Remember to include explanations of key themes, concepts, and decisions taken.

**You do not need to include a cover sheet with part 1, but make sure to include your Student Reference Number (SRN) in the name of the file. Your cover sheet for part 2 will cover part 1 and part 2 of the submission.**

**Ensure you train and evaluate your model before submitting as the examiner will not train every model received.**

## Part 2: Machine Learning and A.I. in the Workplace (50%)

This part of the assessment covers:

- LO1:** Critically evaluate machine learning algorithms and their applicability to business problems

In this section you will develop a proposed machine learning application in the context of your own workplace.

In preparation for this part of the assessment ensure you:

- Identify a potential application for a machine learning and/or A.I. solution in your workplace. This could cover any number of possible solutions such as chatbots, anomaly detection, robotic process automation, character recognition, recommendation systems etc.
- Conduct the required internal research:
  - Engage with your employer on your proposed idea - a good starting point is a discussion with your line manager
  - Identify key stakeholders and develop a stakeholder engagement plan
  - Canvas opinion from potential users and those that are likely to benefit / be impacted by your proposed solution.
- Provide research of the tools you propose to use in developing your solution:
  - This could be a new build, or
  - A cloud-based service provider
- Be sure to provide details of:
  - The machine learning architecture you seek to deploy, ensuring you include all high-level technical specifications such as modules, libraries, and components that you would need to implement your solution such as platforms, databases, interfaces etc.
- Ensure the scope of your proposal is manageable, achievable, and clearly articulates your focus.
- Be sure to view this from a work context and consider your unique environment
- Consider the formative work you have already completed as this should inform your summative assessment

### Part 2 Formative Submission

Your formative submission for Part 2 should be:

- A deck of no more than 14 presentation slides (PowerPoint / Google Slides) making a case for your proposal to the EXCO or appropriate Senior Management Committee.

The suggested format for the slide deck for the (summative) presentation is the following:

- Title slide
- Up to 2 slides – context/background/problem statement
- Up to 2 slides – proposed approach

- Up to 2 slides – theory
- Up to 2 slides – proposed methodology/implementation
- One slide – cost analysis
- One slide – ethics/risks/mitigations
- One slide – impact and evaluation (how will we know the project has worked?)
- One slide – conclusion/recommendation
- One slide – reference list

For the formative submission, it is recommended to provide at least the slides covering the background and proposed approach.

Note – you should avoid text-heavy slides. In order to help the marker provide helpful feedback, you are allowed to use the Power Point “notes” space to complement the slides with some talking points.

## Part 2 Summative Submission

### Submission

Building on the Formative Assessment submission and feedback received, develop your proposal in more detail and develop a recording of **no more than ten minutes** and a deck of **no more than fourteen presentation slides** (PowerPoint / Google Slides, count including title slide and reference slide) making a case for your proposal.

- Your recording must be submitted either as an .mp4 file or a private YouTube link.
  - Please do not use links to private file shares such as One Drive.
- You must supply a separate copy of your slides
- To make the presentation more realistic, you can include a short Q&A session at the end (still included in the ten minutes) where you can answer a maximum of two likely questions from the audience.

In this submission we are looking for further development of your initial idea in your accompanying narrative. The solution should be relevant to your environment and ideally provides your organisation with a realistic commercial opportunity.

Remember this is a recommendation to senior management, some of whom may not be as data fluent as you are; placing a strong emphasis on your ability to communicate effectively. The use of corporate colour palettes and themes is encouraged.

### Research and referencing

You should use your research to strengthen the case for your proposed idea to enable management to take an informed decision to allocate resources to this project. Therefore, ensure you support your proposal with sound academic research and relevant case studies of similar applications in your industry or domain. All references must be listed in the final slide as per suggested template. Any references must be cited in the main body of the slides.

Your research effort should include both:

- Internal research
- External research
  - Using the academic literature to support your points
  - Providing relevant case studies and examples from industry.

You may only use the *Harvard Referencing System* as highlighted in the *General Assessment Guidance* section of this document. All the figures/tables used in the report must have captions and, wherever needed, properly referenced, cited, and explained in your submission.

### **Marking Guide/ Assessment Criteria**

The assignment is 100% of the module grade, with 50% counting towards the Part 1 and 50% counting towards the presentation (Part 2). The marking criteria for both parts are below. Parts one and two will have separate links on the Hub.

### **Submission format**

Please submit a zipped folder including

- BPP Coursework Cover Sheet (covering part 1 and 2)
- The presentation slides
- Recording

<b>Part 1: Machine Learning Models (50% overall)</b>		<b>0-39% Fail</b>	<b>40-49% Low Fail</b>	<b>50-59% Pass</b>	<b>60-69% Merit</b>	<b>70-79% Distinction</b>	<b>80-100% Excellent</b>
20%	1. Import, Prepare and Shape Input Data and EDA	Failure to import, pre-process and shape data for use in machine learning models.	Data is correctly imported but is not appropriately pre-processed negative impacting on potential model outputs	Data imported, processed, and shaped correctly with minimal explanation. Notebook provides a summary of the data and includes descriptive statistics as well as visuals.	Data imported, processed, and shaped correctly with good explanations of the steps taken. Notebook provides a summary of the data and includes descriptive statistics as well as relevant visuals. Student explains in little detail the steps taken to reach the results.	Data imported, processed, and shaped correctly with very good explanations of the steps taken. Notebook provides a summary of the data and includes descriptive statistics as well as relevant visuals. Student explains clearly the steps taken to reach the results.	Data imported, processed, and shaped correctly with very good explanations of the steps taken. Shows high levels understanding of the importance of machine learning data preparation. Notebook provides a summary of the data and includes descriptive statistics as well as relevant visuals developed to a professional standard. Student explains in great detail the steps taken to reach the results.
30%	2. Train and Evaluate Model, Explain Model Architecture.	No model selected / explained. Inappropriate model selected (such as linear model).	Models selected but no accompanying explanation provided.	Correct models provided with basic commentary of model architecture, hyperparameters and evaluation.	Correct models provided with more detailed commentary of model architecture, hyperparameters and evaluation. Some discussion of how hyperparameters were selected and tuned. Models are presented with outputs and clear explanations.	Correct models provided with clear, succinct commentary of model architecture, hyperparameters and evaluation. Clear evidence of understanding hyperparameter selection and model tuning. Models are presented with outputs and with good explanations and insights.	Correct models provided with clear, succinct commentary of model architecture, hyperparameters and evaluation. Clear evidence of understanding hyperparameter selection and model tuning. Models are presented with outputs but with excellent explanation and insights, clearly showing in depth understanding of why selections are made
20%	3. Research	No research conducted or inappropriate methods selected (e.g., irrelevant classification methods or pre-trained models).	Basic research conducted with minimal implementation. Methods selected but no accompanying explanation provided.	Correct methods selected with basic commentary on the research process, implementation, and initial results.	Detailed research conducted with more comprehensive commentary on the research process, implementation, and evaluation. Some discussion of how methods were selected and tuned. Methods are presented with outputs and clear explanations.	Clear, succinct commentary on the research process, implementation, and evaluation. Clear evidence of understanding method selection and tuning. Methods are presented with outputs and good explanations and insights.	Clear, succinct commentary on the research process, implementation, and evaluation. Clear evidence of understanding method selection and tuning. Methods are presented with outputs and excellent explanations and insights, clearly showing in-depth understanding of why selections were made
20%	4. Communicate your Findings	Vague or irrelevant findings such as general explanations provided.	Incorrect interpretation of the model outputs and limitations.	Correctly highlights model outputs and limitations with limited discussion.	Good explanations of how models work, the limitations thereof and steps taken to generate the output.	Very good explanations of how models work, the model limitations, supported and steps taken to generate the output, some support provided in the literature.	Excellent explanations of how models work, the model limitations, supported with references to examples in the literature.
10%	5. References and Support for Key Decisions Taken	No or little effort made to support the arguments made.	Some research provided however not entirely relevant to the arguments and decisions taken.	Research provided is of a general nature.	Good research effort, supporting key decisions.	Very good research effort to support key decisions taken; examples of similar applications provided and primary literature included.	Excellent primary research, case study examples and relevant references provided

Part 2: Machine Learning Proposal for the workplace (50% overall)		0-30-39% Fail	40-49% Low Fail	50-59% Pass	60-69% Merit	70-79% Distinction	80-100% Excellent	
60%	<b>Knowledge base and communication</b>	<b>Inadequate</b> and often implicit knowledge base with some omissions and/or lack of theory relating to the use of machine learning and AI in a work based context. The presentation lacks clarity and coherence. Concepts are poorly organized and difficult to follow.	The proposal shows minimal understanding of ML/AI theory and lacks connection to ethical or workplace issues. Dataset is not described or inappropriate for the proposed method. No clear example or context is provided. Method is unclear or unjustified. Presentation is confusing or disorganized, with poor or missing visual support.	The proposal touches on ML/AI theory, but lacks analysis or clear connection to the workplace. Dataset is mentioned, but not described in sufficient detail. Example or context is vague or underdeveloped. Understanding of the method is partial, with limited justification. Presentation is somewhat organized, but may lack clarity or flow; visuals are present but underused.	The proposal covers key ML/AI concepts, though some areas may lack depth or clarity. Dataset description is present but limited, with minimal discussion of relevance or limitations. Includes a basic example or reference to workplace context, though not fully developed. Shows basic understanding of the proposed method, with some justification. Presentation is structured and understandable, with visuals that aid comprehension.	The proposal clearly explains ML/AI concepts and ethical issues, with some depth and originality. Describes the dataset(s) used, including basic structure and relevance to the problem. Includes a workplace-related example or scenario to contextualize the proposal. Demonstrates a sound understanding of the chosen approach, with appropriate justification. Presentation is clear and coherent, with visual aids that support key points and are accessible to a virtual audience.	The proposal deeply analyses ML/AI theory and ethical considerations, linking them to the workplace context. Provides detailed dataset descriptions, including rationale for selection and potential challenges. Includes at least one relevant case study or example to support the proposal. Shows independent thinking and a strong grasp of appropriate methods and tools. Presentation is engaging and well-structured, with visuals that support key points and are accessible to a virtual audience.	The proposal critically explores machine learning and AI theory, including advanced concepts, ethical implications, and real-world relevance. Includes clear descriptions of datasets, including source, structure, limitations, and suitability for the proposed ML/AI approach. Uses multiple relevant case studies or examples to illustrate application in a work-based context. Demonstrates original thinking, innovation, and autonomy in selecting and justifying methods. Communication is clear, confident, and well-paced, with visual aids that enhance understanding and are tailored for online delivery.
30%	<b>Research skills (Internal and External)</b>	<b>Inadequate</b> critical analysis or evaluation with some difficulties. Largely imitative and descriptive. Some difficulty with structuring the line of logical argument and accuracy in expression of argument.	Little or no research is used to support the proposal. Sources are irrelevant, outdated, or missing entirely. No attempt to analyse or critique information.	<b>L</b> <b>Uses few or inappropriate sources, such as blogs or unverified websites.</b> <b>Research is superficial, with little connection to the proposal's aims.</b> <b>No evidence of critical thinking or comparison between sources.</b>	Uses a limited number of sources, with some relevance to the topic. Research is mostly descriptive, with limited critical analysis. Some sources may be outdated or not clearly linked to the proposal.	Uses a range of relevant sources, mostly appropriate and current. Some evidence of critical engagement, though analysis may be uneven. Research supports the proposal but may lack depth in some areas.	Uses a strong mix of academic and professional sources to support key aspects of the proposal. Shows critical thinking, such as questioning assumptions or comparing methods. Research is clearly linked to the workplace context and project goals.	Demonstrates extensive and critical engagement with a wide range of sources, including: Peer-reviewed academic literature Industry reports Technical documentation Ethical frameworks Research is used to frame the problem, justify the approach, and evaluate alternatives. Sources are compared and critiqued, showing awareness of limitations, biases, and differing perspectives. Includes recent and relevant publications, showing up-to-date knowledge of ML/AI developments.

10%	<i>Referencing</i>	Referencing is absent.	Referencing is highly inaccurate	Many sources are <b>not explicitly referenced</b>	<i>Referencing is mostly correct, but may include some inconsistencies. Most sources are cited, though some may be referenced only implicitly.</i>	Referencing is mostly accurate and consistent, with minor errors. All key sources are clearly cited, and a reference list is provided.	<b>All sources are explicitly cited using a consistent and correct referencing format (Harvard). Referencing is seamlessly integrated into the presentation.</b> Includes in-text citations and a complete reference list.	All sources are explicitly cited using a consistent and correct referencing format (Harvard). Referencing is seamlessly integrated into the presentation. Includes in-text citations and a complete reference list.
-----	--------------------	------------------------	----------------------------------	---	--	---	---	--

## General Assessment Guidance

- You are required to submit all elements of your assessment via **Turnitin online access**. Only submissions made via the specified mode will be accepted and hard copies or any other digital form of submissions (like via email or pen drive etc.) **will not be accepted**.
- **Please ensure your student registration number is on your front cover sheet**
- You are required to achieve minimum **50%** to **pass** this module and must address all learning outcomes.
- You are required to use **only Harvard Referencing System** in your submission. Any content which is already published by other author(s) and is not referenced will be considered as a case of plagiarism.
- You can find further information on Harvard Referencing in the online library on the Hub. You can use the following link to access this information:  
<https://bpp.libguides.com/c.php?g=699362&p=5022940>
- BPP University has a strict policy regarding authenticity of assessments. In proven instances of plagiarism or collusion students will go through the malpractice process. You are advised to read the rules and regulations regarding plagiarism and collusion in the **BPP University Policies and Procedures**
- BPP University have published guidelines on the **Acceptable Use of GenerativeAI** which we expect students to familiarise themselves with. In addition, you can find a wealth of information on the Hub [here](#), under BPP's 'Generative AI Awareness' module. You will also find a tool '[Can I use AI in my studies at BPP?](#)' to help you determine whether your intended use of Generative AI is allowed in a particular type of programme or course.
- You **should include** a completed copy of the **Assignment Cover sheet**. Any submission **without** this completed Assignment Cover sheet may be considered **invalid and not marked**.
- Please note that any work submitted in an appendices is for information only and is NOT marked. You should ensure you hit all requirements and learning outcomes within the word count and time set.
- Assessment briefs contain strict and specific requirements on word counts and word count equivalents, such as slide number or presentation duration. These are not recommendations, they are formal and assessed requirements of the brief. You may submit work below the word count but you may not exceed the word count or word count equivalent.

Should you exceed the word count or allotted time your assessor will stop marking at the point where the word count has been reached. Any work following this point will not contribute to your grade. This applies to both formative and summative submission.