AFRICA UNIVERSITY
A United Methodist-Related Institution

"Investing in Africa's Future"

COLLEGE OF ENGINEERING AND APPLIED SCIENCES

NCSC114: PRINCIPLES OF PROGRAMMING LANGUAGES
NHAI114 FUNDAMENTALS OF PROGRAMMING FOR MACHINE LEARNING

END OF FIRST SEMESTER EXAMINATIONS

NOVEMBER 2024

LECTURER: MR BRAITON U MUKHALELA

TIME: 3 HOURS

INSTRUCTIONS

Create a folder on the desktop of your allocated computer into which you are to save all your work. The folder name should be your Student Reg Number- Course code eg **20001-NCIS301**

You are required to answer ALL questions as instructed in each section.

Use Invigilator Allocated Computer in one of the designated Computer Labs

The Examination is Lab based, no any other writing material is required.

Section A: General Python Programming (1 hour) (Choose either

Question 1

You have been asked to design a Python based SQL storage App (database) for a newly established AU-Hospital. The App should have the following entities for storage of data: Ward, Doctor, Patient, Drug, and Prescription.

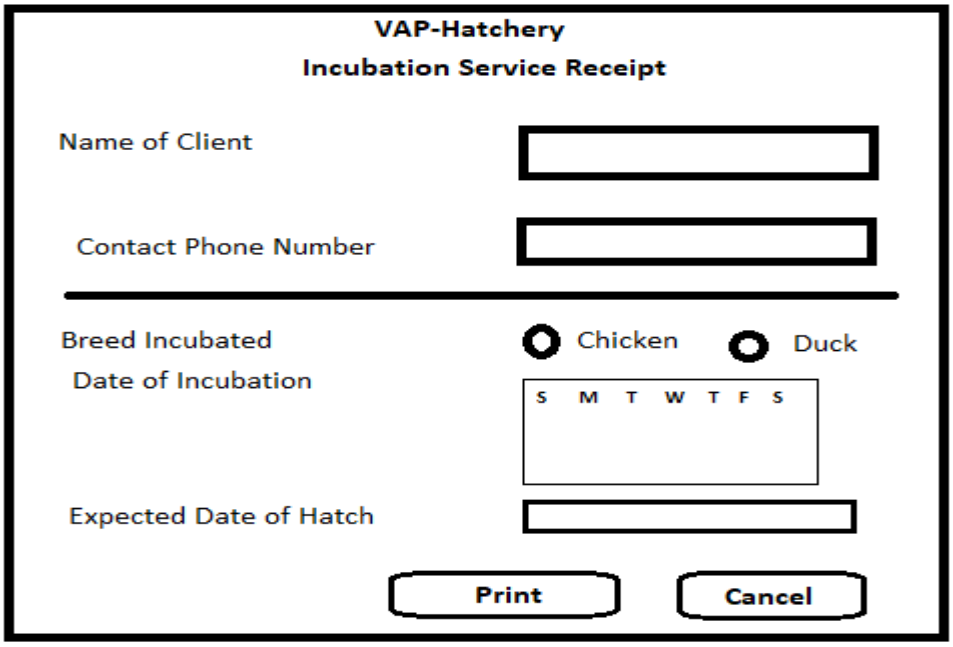
Required: Program the App using Python and embed SQL statement for doing the following queries:

i) Adding at least 3 records for each of the identified entities.
ii) Show all patients who have at one time been treated by a certain Doctor.
iii) All Patients who have at one time been administered a certain drug. [20]

Question 2

(a) Create a story to be included on a Random Story Teller App by a Kid Cartoons developer. NB: Make use of lists, conditional statements, random function etc. [8]

i. Your Senior Font End developer has given you the scratch below for a client into Hatchery who wants the proposed GUI shown below. Write a python program to produce a professional application with an orange background. Consider the following facts in Hatchery business to be embed into the algorithms: Chicken eggs take 21 days to Hatch and Duck eggs take 28 days to hatch, thus say Date of incubation is today, the system should thus show the Expected Date of Hatch in the appropriate field, depending on the spices of eggs incubated ie Chicken and Duck for now.



[12]

Section 2 : Programming for Machine Learning

NB: Question 3 is Compulsory for all CS and SE majors and choose any other 2 questions, whereas for AI majors Choose any 3 as from Question 4)

Question 3 (30 minutes)

a) You have just joined a mining company Vusu mines and your first task was to write a python program to be embed into an Ardiuno Uno Controller that is to control a wind turbines safety system. You are told that a wind turbine has a safety system which uses three inputs to a logic circuit. A certain combination of conditions results in an output , X, from the logic circuit being equal to 1. When the value of X = 1, the wind turbine is shut down. The following table shows which parameters are being monitored and form the three inputs to the logic circuit.

Parameter Description	Parameter	Binary Value	Description of Condition
Turbine Speed	S	0	Turbine speed < or = 1000rpm
		1	Turbine speed > 1000rpm
Bearing temperature	T	0	Bearing temperature < or = 80 degrees Celsius
		1	Bearing temperature > 80 degrees Celsius
Wind velocity	W	0	Wind velocity < or = 120 kph

			Wind velocity > 120 kph

[20]

Question 4: Data Manipulation with Pandas (30 minutes)

Datasets: Iris Dataset

You can use the classic Iris dataset for various tasks. It can it has been included for your convenience in a folder called **Programming for ML workfiles** or imported directly using libraries like `seaborn`.

1. **Loading Data**
   - Load the Iris dataset using Pandas. Display the first five rows of the dataset.
2. **Data Cleaning**
   - Check for any missing values in the dataset and handle them appropriately (if applicable).
3. **Data Analysis**
   - Calculate the average sepal length for each species of the iris flower. **(20 marks)**

Question 5: Data Visualization (30 minutes)

1. **Data Visualization**
   - Use Matplotlib and Seaborn to create the following:
     - A pairplot of the Iris dataset to visualize the relationships between features.
     - A boxplot showing the distribution of sepal lengths for each species. **(20 marks)**

Question 6: Supervised Learning (30 minutes)

1. **Model Building**
   - Split the dataset into training and testing sets (80% train, 20% test).
   - Implement a decision tree classifier using `scikit-learn` to classify the species of iris based on the features.
   - Fit the model to the training data and evaluate its accuracy on the test data.
2. **Model Evaluation**
   - Generate a confusion matrix and classification report for the model.
   - Discuss the implications of overfitting and underfitting in the context of your results. **(20 marks)**

Question 7: Hyperparameter Tuning (30 minutes)

1. **Hyperparameter Tuning**
    - Use GridSearchCV to tune the hyperparameters of your decision tree model. Focus on parameters such as `max_depth` and `min_samples_split`.
    - Report the best parameters and the corresponding accuracy. **(20 marks)**

Submission Guidelines

- Submit your Python scripts (.py files).
- Ensure all code is well-commented for clarity.
- Include visualizations as part of the notebook or as separate image files.

Evaluation Criteria

- **Correctness of Solutions:** 50%
- **Code Quality and Clarity:** 30%
- **Data Visualization Quality:** 20%

END OF EXAMINATION