****

Faculty of Computers and Artificial Intelligence

Computer Science Department

2021/2022

**CS 395 Selected Topics in CS-1**

**Research Project**

Report Submitted for Fulfillment of the Requirements and ILO’s for Selected Topics in CS-1 course for Fall 2021

Team No. 44

|  |  |  |  |
| --- | --- | --- | --- |
|  | ID | Name | Grade |
|  | 201900144 | اسلام محمد فايز عبد المطلب |  |
|  | 201900200 | اياد عبد المنعم عبد الفتاح احمد |  |
|  | 201900197 | أنور محمد أنور حسن |  |
|  | 201900936 | هدير عماد الدين جاد محمد |  |
|  | 201900917 | نورهان وليد محمد عبد العظيم |  |
|  | 201900799 | مريم اسامه محمد احمد انيس |  |
|  | 201900373 | شمس طه عبد العزيز القاضي |  |

Delivered to:

**Dr. Wessam El-Behaidy**

**Eng. Islam Gamal**

**Eng. Muhammed Kamal**

I. NUMERICAL DATASET

1. Project Introduction

* 1. **Dataset Name**

online\_shoppers\_intention‏

**https://www.kaggle.com/roshansharma/online-shoppers-intention**

* 1. **Number of classes and their labels**

17 class & 1 label

* 1. **Dataset Samples Numbers**

12330

* 1. **Training, Validation and Testing**

**Training (60%): 7891**

**Validation (20%): 1973**

**Testing (20%):2466**

|  |  |  |
| --- | --- | --- |
| **Training**  **80%**  **9864 sample** | | **Testing**  **20%**  **2466 sample** |
| **Training**  **60%**  **7891 sample** | **Validation**  **20%**  **1973 sample** |

1. Implementation Details
   * 1. **Extracted Features**

(How many features were extracted, their names, the dimension of resulted features)

11 Features

Names: - Administrative , Administrative\_Duration Informational , Informational\_Duration , ProductRelated BounceRates , PageValues , SpecialDay , Weekend Visitor\_New\_Visitor , Visitor\_Returning\_Visitor

the dimension of resulted features: 12330

* + 1. **Cross-validation**

(Is cross-validation is used in any of implemented models? If yes, specify the number of fold and ratio of training/validation)

**NO need cross validation because dataset is big**

* + 1. **Artificial Neural Network (ANN)**
* **Hyper-parameters**

(Specify all the hyper-parameters (initial learning rate, optimizer, regularization, batch size, no. of epochs…) with their specified value in implementation)

Initial Learning: - 0.001

Optimizer: - Adam (adaptive optimizer)

Regularization : No need Regularization because no overfitting

Batch Size: 400

NO.OF.EPHOCS: 100

* + 1. **Support Vector Machine** **(SVM)**
* **Hyper-parameters**

(Specify all the hyper-parameters (optimizer, regularization, …) with their specified value in implementation)

C=1

Kernel = linear

Regularization : no regularization as there is no overfitting

Gamma : default value (scale)

1. Models Results

**For each model you should show all these results for your model on testing data** (loss curve, accuracy, confusion matrix, ROC curve)

* 1. **ANN Results**

**Loss Curve**

**Chart

Description automatically generated**

**Accuracy & confusion matrix**

**Text

Description automatically generated**

**ROC curve**

**Chart

Description automatically generated with medium confidence**

* 1. **SVM Results**

**Confusion Matrix**

Accuracy : 88.07%

Confusion matrix :

[[1860 214]

[ 80 312]]

Learning Curve

Chart, line chart

Description automatically generated

ROC

Chart, line chart

Description automatically generated

II. IMAGE DATASET

1. Project Introduction

* 1. **Dataset Name**

Fashion -minist

Trouser - T shirt

* 1. **Number of classes and their labels**

2 classes

Trouser

Tshirt

* 1. **Dataset Images Numbers and size**

14000

* 1. **Training, Validation and Testing**

**Training: 12000**

**Testing: 2000**

|  |  |
| --- | --- |
| **Training**  **72%**  **12000 sample** | **Testing**  **28%**  **2000 sample** |

2. Implementation Details

* + 1. **Extracted Features**

**Image (28\*28) and make image black White**

* + 1. **Cross-validation**

**We do not use cross validation**

* + 1. **Artificial Neural Network (ANN)**
* **Hyper-parameters**

(Specify all the hyper-parameters (initial learning rate, optimizer, regularization, batch size, no. of epochs…) with their specified value in implementation)

Initial learning rate:- 0.001

Optimizer : adam

No batch sizes

Epochs :- 10

* + 1. **Support Vector Machine** **(SVM)**
* **Hyper-parameters**

(Specify all the hyper-parameters (optimizer, regularization, …) with their specified value in implementation)

C=1

Kernel = poly

Regularization : no regularization as there is no overfitting

Gamma : auto

3. Models Results

**For each model you should show all these results for your model on testing data** (loss curve, accuracy, confusion matrix, ROC curve)

* 1. **ANN Results**

**Loss curve**

**Chart, line chart

Description automatically generated**

**Roc**

**Chart, line chart

Description automatically generated**

**Confusion matrix**

**A picture containing text

Description automatically generated**

* 1. **SVM Results**

**ROC Curve**

**Graphical user interface

Description automatically generated**

**Confusion Matrix**

**A picture containing text

Description automatically generated**