



Co-funded by the European Union



Norwegian Embassy
Islamabad



ARTIFICIAL INTELLIGENCE DATA TECHNICIAN



© TVET SSP

CBT CURRICULUM

National Vocational Certificate Level 4

Version 1 - November, 2019



Implemented by

giz Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

Published by

National Vocational and Technical Training Commission
Government of Pakistan

Headquarter

Plot 38, Kirthar Road, Sector H-9/4, Islamabad, Pakistan
www.navttc.org

Responsible

Director General Skills Standard and Curricula, National Vocational and Technical Training Commission
National Deputy Head, TVET Sector Support Programme, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Layout & design

SAP Communications

Photo Credits

TVET Sector Support Programme

URL links

Responsibility for the content of external websites linked in this publication always lies with their respective publishers. TVET Sector Support Programme expressly dissociates itself from such content.

This document has been produced with the technical assistance of the TVET Sector Support Programme, which is funded by the European Union, the Federal Republic of Germany and the Royal Norwegian Embassy and has been commissioned by the German Federal Ministry for Economic Cooperation and Development (BMZ). The Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH in close collaboration with the National Vocational and Technical Training Commission (NAVTTTC) as well as provincial Technical Education and Vocational Training Authorities (TEVTAs), Punjab Vocational Training Council (PVTC), Qualification Awarding Bodies (QABs)s and private sector organizations.

Document Version

November, 2019
Islamabad, Pakistan

ARTIFICIAL INTELLIGENCE DATA TECHNICIAN



© TVET SSP

CBT CURRICULUM
National Vocational Certificate Level 4

Version 1 - November, 2019

Introduction	4
Definition/ Description of the training program for Artificial Intelligence Data Technician	4
Purpose of the training program	4
Overall objectives of training program	4
Competencies to be gained after completion of course	6
Possible available job opportunities available immediately and later in the future	7
Trainee entry level	7
Minimum qualification for trainer	7
Recommended trainer: trainee ratio	7
Medium of instruction i.e. language of instruction	7
Duration of the course (Total time, Theory & Practical time)	7
Sequence of the modules	8
Summary – Overview of the curriculum	9
Modules	9
Module: Scrape data from the web (061900931)	9
Module: Process Images through Image Processing software (061900932)	16
Module: Work with Data Manipulation Toolkit (061900933)	20
Module: Work with Multidimensional Arrays' Manipulation and Computation Package (061900934)	25
General assessment guidance for <i>Artificial Intelligence Data Technician</i>	33
Complete list of tools and equipment – Non Consumable	36
Complete list of tools and equipment - Consumable	36
Credit values	38

Introduction

Definition/ Description of the training program for Artificial Intelligence Data Technician

In order to build the capacity of technical and vocational training institutes in Pakistan through provision of demand driven competencies-based trainings in Information technology sector the NAVTTC, and TEVT Sector Support Program (TSSP) have joined hands together to develop qualifications for Information Technology sector. These qualifications will not only build the capacity of existing workers of this sector but also support the youth to acquire skills best fit for this sector. The benefits and impact of development of these qualifications will be on both demand and supply side.

Based upon this demand of industry these competency-based qualifications for Artificial Intelligence Data Technician are developed under National Vocational Qualification Framework (NVQF) (Level 1 to 4). The qualifications mainly cover competencies along with related knowledge and professional skills which are essential for getting a job or self-employed.

The qualifications are also in line with the vision of Pakistan's National Skills Strategy (NSS), National TVET Policy and National Vocational Qualification Framework (NVQF). This provides policy directions, support and an enabling environment to the public and private sectors to impart training for skills development to enhance social and economic profile. The National Vocational & Technical Training Commission (NAVTTC) has approved the Qualification Development Committee (QDC). The QDC consists experts from the relevant industries from different geographical locations across Pakistan and academicians who were consulted during the development process to ensure input and ownership of all the stakeholders. The National Competency Standards could be used as a referral document for the development of curricula to be used by training institutions.

Purpose of the training program

The purpose of the training is to provide skilled manpower to improve the existing capacity of Information Technology sector. This training will provide the requisite skills to the trainees to become Artificial Intelligence Data Technician. It will enable the participants to meet the challenges in the field of Artificial Intelligence. Further, to improve the skill level of the technician and prepare them for the information technology industry to meet the market competition nationally and internationally.

The core purpose of this qualification is to produce employable Artificial Intelligence Data Technicians who can work as Artificial Intelligence Data Technician according to national and international standards. In addition, this qualification will prepare unemployable youth to employees in this sector.

Overall objectives of training program

The Artificial Intelligence Data Technician qualification from level 1- 4 consists of theoretical and practical details required for Artificial Intelligence Data Technician in information technology industries. However, this will require providing additional input on

entrepreneurship development for the one who is willing to start his/her own business. The main objective of the qualification is to prepare Artificial Intelligence Data technician having set of skills as follows:

- Comply with Work Health and Safety Policies
- Obey the Workplace Policies and Procedures
- Follow Basic Communication Skills (General)
- Demonstrate Basic Literacy Skills
- Operate Computer Functions (General)
- Use Word Processing Software
- Use of Spreadsheet
- Comply Personal Health and Safety Guidelines
- Communicate the Workplace Policy and Procedure
- Perform Basic Communication (Specific)
- Demonstrate Basic Numeracy Skills
- Use Multimedia Processing
- Pre-Process Data
- Perform Basic Computer Application (Specific)
- Apply Work Health and Safety Practices (WHS)
- Identify and Implement Workplace Policy and Procedures
- Communicate at Workplace
- Manage Personal Finances
- Code in Programming Language suitable for AI
- Setup Environment
- Perform Computer Application Skills
- Contribute to Work Related Health and Safety (WHS) Initiatives
- Comply with Workplace Policy and Procedures
- Perform Advanced Communication
- Manage Human Resource Services
- Scrape data from the web
- Process Images through Image Processing software
- Work with Data Manipulation Toolkit
- Work with Multidimensional Arrays' Manipulation and Computation Package
- Develop Advance Computer Application Skills
- Develop Entrepreneurial Skills

Competencies to be gained after completion of course

At the end of the course, the trainee must have attained the following competencies:

1. Comply with Work Health and Safety Policies
2. Obey the Workplace Policies and Procedures
3. Follow Basic Communication Skills (General)
4. Demonstrate Basic Literacy Skills
5. Operate Computer Functions (General)
6. Use Word Processing Software
7. Use of Spreadsheet
8. Comply Personal Health and Safety Guidelines
9. Communicate the Workplace Policy and Procedure
10. Perform Basic Communication (Specific)
11. Demonstrate Basic Numeracy Skills
12. Use Multimedia Processing
13. Pre-Process Data
14. Perform Basic Computer Application (Specific)
15. Apply Work Health and Safety Practices (WHS)
16. Identify and Implement Workplace Policy and Procedures
17. Communicate at Workplace
18. Manage Personal Finances
19. Code in Programming Language suitable for AI
20. Setup Environment
21. Perform Computer Application Skills
22. Contribute to Work Related Health and Safety (WHS) Initiatives
23. Comply with Workplace Policy and Procedures
24. Perform Advanced Communication
25. Manage Human Resource Services
26. Scrape data from the web
27. Process Images through Image Processing software
28. Work with Data Manipulation Toolkit
29. Work with Multidimensional Arrays' Manipulation and Computation Package
30. Develop Advance Computer Application Skills
31. Develop Entrepreneurial Skills

Possible available job opportunities available immediately and later in the future

Artificial Intelligence Data Technician are employed in Information Technology Sector. Experienced Artificial Intelligence Data Technician may advance through promotions with the same employer or by moving to more advanced positions with other employers. They can become:

- Artificial Intelligence Data Technician
- Artificial Intelligence Data Supervisor
- Artificial Intelligence Data Trainer
- Stock Exchange Predictor
- Artificial Intelligence Big Data Analyst
- Data Entry Operator
- Financial Forecasting
- Media House Data Technician

Trainee entry level

- Middle (Grade 8) for level-1
- Level-1 for level-2
- Level-2 for level-3
- Level-3 for level-4

Minimum qualification for trainer

- BS in (Artificial Intelligence/Data Science/Computer Science/Computer Engineering/Software Engineering/Information Technology/Electrical/Mechatronics) or relevant fields.

Recommended trainer: trainee ratio

The recommended maximum trainer: trainee ratio for this program is 1 trainer for 20 trainees

Medium of instruction i.e. language of instruction

Instructions will be in English/Urdu language.

Duration of the course (Total time, Theory & Practical time)

This curriculum comprises of 10 modules. The recommended delivery time for technical units is 550 hours.

- Delivery of the course can therefore be full time (8 hours a business day), 6 days a week, for 24 months (on average 26 working days a month) for each level. Training providers are at liberty to develop other models of delivery, including part-time and evening delivery. OR
- Delivery of the course can therefore be full time (9 hours a business day), 5 days a week, for 24 months (on average 22 working days a month). Training providers are at liberty to develop other models of delivery, including part-time and evening delivery.

The full structure of the course is as follows:

Sr. No	Module	Theory hours	Workplace hours	Total hours
01	Contribute to Work Related Health and Safety (WHS) Initiatives			
02	Comply with Workplace Policy and Procedures			
03	Perform Advanced Communication			
04	Manage Human Resource Services			
05	Scrape data from the web	20	110	130
06	Process Images through Image Processing software	50	70	120
07	Work with Data Manipulation Toolkit	30	120	150
08	Work with Multidimensional Arrays' Manipulation and Computation Package	30	120	150
09	Develop Advance Computer Application Skills			
10	Develop Entrepreneurial Skills			

Sequence of the modules

This qualification is made up of 10 modules. A suggested distribution of these modules is presented overleaf. This is not prescriptive and training providers may modify this if they wish.

The following technical module will be followed as require for the training purpose.

Module 05

Module 06

Module 07

Module 08

Each module covers a range of learning components. These are intended to provide detailed guidance to teachers (for example the Learning Elements component) and give them additional support for preparing their lessons (for example the Materials Required component). The detail provided by each module will contribute to a standardized approach to teaching, ensuring that training providers in different parts

of the country have clear information on what should be taught. Each module also incorporates the industrial demand of Pakistan that make this qualification unique to Pakistan's industry needs.

Summary – Overview of the curriculum

At the completion of this curriculum candidate will be able to contribute and follow to health safety. Candidate will be able to process image, use different data manipulation tool kits and work with multi-dimensional array's.

Modules

Module: 061900931 Scrape data from the web

Objective of the Module: After the completion of this module the candidate will be able to Scrape different types of data from the Web, build web pages and extract data from a variety of e-sources.

Duration: 130 hrs. Theory: 20 hrs. Practice: 110 hrs.

Learning Unit	Learning Outcomes	Learning Elements	Duration	Material/Tools Required	Learning Place
LU1: Build a static web page	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Implement basic HTML tags 2. Implement basic HTML attributes usage. 3. Implement basic JavaScript behaviors. 4. Perform inspection of a webpage. 	<ul style="list-style-type: none"> • Describe basics of Internet and the World Wide Web • Demonstrate basic HTML tags • Demonstrate basic attributes of HTML tags • Demonstrate syntax of HTML document • Describe and demonstrate basic syntax of JavaScript • Demonstrate basic functionality of web browser 	<p>Total 24 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 20 Hrs</p>	<ul style="list-style-type: none"> • Computer system • Internet facility • Python • BeautifulSoup • Request • Drivers 	<p>Theory: Class</p> <p>Practical Lab</p>

	<p>5. Create a basic webpage</p>	<ul style="list-style-type: none"> • Demonstrate understanding of developer tools in web browsers • Demonstrate understanding of basic web development tools • Demonstrate understanding of web process development 			
		<p>Practical-1 View a simple webpage and explore its structure using a web browser's developer tools.</p> <p>Practical-2 Create a website consisting of multiple static webpages using basic HTML tags or Java Script view it in a browser</p>			
<p>LU2: Extract data as per requirement</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Set request headers. 2. Set request cookie values where required 3. Configure a driver to some browser as required 4. Generate a request to webserver 5. Load response stream 6. Convert stream to page source/content 	<ul style="list-style-type: none"> • Perform installation requests library for python • Demonstrate basics of Internet and the World Wide Web • Demonstrate basic HTML tags • Demonstrate basic attributes of HTML tags • Demonstrate syntax of HTML document • Describe and demonstrate basic syntax of JavaScript • Demonstrate basic functionality of web browser 	<p>Total 28 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 24 Hrs</p>		<p>Theory: Class</p> <p>Practical Lab</p>

	<p>7. Read response headers</p>	<ul style="list-style-type: none"> • Demonstrate understanding of developer tools in web browsers • Demonstrate understanding of basic web development tools • Demonstrate understanding of web process development basic functionality of requests module • Describe methods of requests module • Describe common request headers • Demonstrate usage of get method in requests module • Describe functionality of HTTP cookie • Demonstrate usage of Cookie and RequestsCookieJar objects in python • Describe basic functionality of web driver for various browsers • Install and configure various web drivers • Demonstrate usage of get method in requests module to generate HTTP GET request • Describe basic functionality of 			
--	---------------------------------	---	--	--	--

		<p>Response object</p> <ul style="list-style-type: none"> • Demonstrate usage of content method in Response object • Demonstrate decoding HTML data • Demonstrate usage of header method in Response object 			
		<p>Practical-1 Develop a python program that downloads a webpage also display headers of webpage and raw data of webpage.</p> <p>Practical-2 Develop a python program that sends cookies to a webserver while requesting a webpage</p>			
<p>LU3: Parse a web page with appropriate package</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Perform installation of beautiful soup 2. Import package into program 3. Request a content to download 4. Find required content from page source 5. Append content 6. Convert content to a data frame 7. Export data 	<ul style="list-style-type: none"> • Describe the basic steps for installing python packages • Describe the basic steps for importing packages into python programs • Demonstrate how to generate web request using Requests package • Describe basic steps to search for specific content using find_all method of beautiful soup method • Demonstrate steps required to 	<p>Total 28 Hrs</p> <p>Theory: 04Hrs</p> <p>Practical: 24 Hrs</p>	<p>Theory: Class</p> <p>Practical Lab</p>	

		<p>append strings to HTML content</p> <ul style="list-style-type: none"> • Describe basic functionality of Dataframe • Demonstrate steps required for extracting tabular data from a webpage and convert to Dataframe • Describe common file formats for storing data • Demonstrate steps required to export data to files using BeautifulSoup package 			
		<p>Practical-1 Develop a python program that can download a webpage and find specific text and append specified content with in the page using BeautifulSoup package</p> <p>Practical-2 Develop a python program that can extract tabular data from a webpage and convert it into Dataframe object. Also export data scraped from the web into CSV/text file format.</p>			
<p>LU4: Extract data from an HTML tag</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Find tag by name 2. Find tag by attribute values 	<ul style="list-style-type: none"> • Describe various find methods available in BeautifulSoup package • Demonstrate how to use the 	<p>Total 24 Hrs</p> <p>Theory: 04 Hrs</p>		<p>Theory: Class</p> <p>Practical Lab</p>

	<p>3. Navigate through values.</p> <p>4. Retrieve tag values</p> <p>5. Retrieve attribute values.</p>	<p>find methods to locate HTML tags in a webpage</p> <ul style="list-style-type: none"> • Demonstrate how to use to find method to locate HTML tags in a webpage have specific values for an attribute • Demonstrate how to iterate through all tag/attribute values • Demonstrate how to get values of tags • Demonstrate how to get values of attributes 	<p>Practical: 20 Hrs</p>		
<p>LU5: Parse xml /JSON</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Read xml/json file. 2. Create xml/json object. 3. Forward navigating 	<ul style="list-style-type: none"> • Describe basic syntax of XML • Describe basic syntax of JSON • Install minidom and ElementTree modules • Demonstrate 	<p>Total 26 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 22 Hrs</p>		<p>Theory: Class</p> <p>Practical Lab</p>

	<p>through elements.</p> <p>4. Backward navigation through elements.</p> <p>5. Navigate through XPath.</p>	<p>how to read XML/JSON files using minidom and ElementTree in python</p> <ul style="list-style-type: none"> • Demonstrate how to create XML/JSON objects in python. • Describe the basic steps of navigating XML/JSON files • Demonstrate the steps required to navigate through XML/JSON file using python • Demonstrate how to use XPath for navigating XML/JSON files in python 			
		<p>Practical-1</p> <p>Develop a python program that can perform following:</p> <ul style="list-style-type: none"> • Read XML/JSON file and XML/JSON file • Display contents of XML/JSON file by navigating the file in forward and reverse. • Navigate and display the contents of XML/JSON file using XPath 			

Module: 061900932 Process Images through Image Processing software

Objective of the Module: This competency standard deals with the skills and knowledge required to perform basic image manipulation/Editing operation.

Duration: 120 hrs. Theory: 50 hrs. Practice: 70 hrs.

Learning Unit	Learning Outcomes	Learning Elements	Duration	Material/Tools Required	Learning Place
LU1: Perform Basic Image Operations	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Read image from file 2. Display an image from data 3. Perform global threshold 4. Perform adaptive thresholding 5. Perform image sharpening 6. Perform image blurring using averaging 7. Perform image blurring using median 8. Perform image blurring using Gaussian 9. Perform image 	<ul style="list-style-type: none"> • Understand the meaning of various image editing operations and image contours <hr/> <p>Practical-1 Using a specified software, read an image, display that image, perform global thresholding and display the image after performing these operations.</p> <p>Practical-2 Perform cropping, adaptive thresholding, image sharpening and blurring -using averaging, median and Gaussian filters- on a given image. And use suitable filters to find image contours</p>	<p>Total 20 Hrs</p> <p>Theory: 10 Hrs</p> <p>Practical: 10 Hrs</p>	<ul style="list-style-type: none"> • Computer system with high performance GPU(s) with graphic card • High speed high capacity storage • High resolution display • Suitable compiler to install and run OpenCV • OpenCV software package, latest version 	<p>Theory: Class</p> <p>Practical: Lab</p>

	cropping 10. Find image contours				
LU2: Apply Image Filters	You will be able to 1. Creating 2D convolution filter 2. Apply Laplacian filter for edge detection 3. Apply X, Y Sobel filter on noisy images 4. Apply canny edge detection filter 5. Plot filtered images	<ul style="list-style-type: none"> Understand the concept of filters in images. <hr/> Practical-1 Given a series of images, apply the assigned filters and display the results.	Total 20 Hrs Theory: 05 Hrs Practical: 15 Hrs		Theory: Class Practical: Lab
LU3: Change Color Spaces	You will be able to 1. Perform RGB to greyscale conversion 2. Perform RGB to HSV conversion 3. Perform RGB to LAB colour conversion 4. Perform RGB to YCrCb color conversion	<ul style="list-style-type: none"> Understand what are image color spaces. <hr/> Practical-1 Apply specified color space conversions using built-in functions.	Total 10 Hrs Theory: 05 Hrs Practical: 05 Hrs		Theory: Class Practical: Lab

LU4: Perform Geometrical Transformation	You will be able to <ol style="list-style-type: none"> 1. Perform scaling operation on image 2. Perform image translation 3. Perform image rotation to any angle 4. Perform affine transformation 	<ul style="list-style-type: none"> • Understand the concept of geometrical transformations <hr/> Practical-1 Apply following transformations on a given image using built-in functions: <ol style="list-style-type: none"> 1. Scaling 2. Translation 3. Rotation Affine transformation	Total 20 Hrs Theory: 10 Hrs Practical: 10 Hrs		Theory: Class Practical: Lab
LU5: Perform Morphological Operations	You will be able to <ol style="list-style-type: none"> 1. Perform image opening 2. Perform image erosion 3. Perform image dilation 4. Perform image closing 5. Perform morphological erosion 6. Perform top hating on image 	<ul style="list-style-type: none"> • Understand the effects and importance of morphological operations. <hr/> Practical-1 Apply following morphological operations on a given image using built-in functions: <ol style="list-style-type: none"> 1. Image opening 2. Erosion 3. Dilation 4. Image closing 5. Morphological erosion 6. Top hating 	Total 20 Hrs Theory: 05 Hrs Practical: 15 Hrs		Theory: Class Practical: Lab
LU6: Match Image	You will be able to	<ul style="list-style-type: none"> • Understand the concept of image 	Total 20 Hrs		Theory: Class

<p>Templates for object Recognition</p>	<ol style="list-style-type: none"> 1. Apply min max lock function 2. Perform template based object matching 3. Perform feature based object matching 4. Perform area based object matching 	<p>templates for object recognition</p> <hr/> <p>Practical-1 Perform min-max lock function and template based object matching</p> <p>Practical-2 Perform feature based and area based object matching</p>	<p>Theory: 10 Hrs</p> <p>Practical: 10 Hrs</p>		<p>Practical: Lab</p>
<p>LU7: Extract Foreground from the Image</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Apply grabcut technique for foreground extraction 2. Prepare image mask of suitable size 3. Apply image mask for foreground extraction 4. Perform series of basic image operations to extract foreground 	<ul style="list-style-type: none"> • Understand the importance of foreground extraction <hr/> <p>Practical-1 Perform foreground extraction using grabcut technique or using series of basic image operations.</p> <p>Practical-2 Prepare and apply image masks for foreground extraction</p>	<p>Total 10 Hrs</p> <p>Theory: 05 Hrs</p> <p>Practical: 05 Hrs</p>		<p>Theory: Class</p> <p>Practical: Lab</p>

Module: 061900933 Work with Data Manipulation Toolkit

Objective of the Module: After the completion of this module the candidate will be able to operate the manipulation toolkit, perform a variety of manipulation operations to import/export the data.

Duration: 150 hrs. Theory: 30 hrs.

Practice: 120 hrs.

Learning Unit	Learning Outcomes	Learning Elements	Duration	Material/Tools Required	Learning Place
LU1: Import Data	You will be able to 1. Open a python script 2. Import pandas 3. Import a csv file using “read_csv” function 4. Import an excel file using “read_excel” function 5. Import from any other file type using appropriate “read” function	<ul style="list-style-type: none"> Importing Pandas library Understand of plain data, columnar data, array data and its operations. Understand data-stream and data-frame 	Total 15 Hrs Theory: 05 Hrs Practical: 10 Hrs	<ul style="list-style-type: none"> Computer system with python installed Pip package manager Stable version of pandas Text/Code editing application Terminal/cmd application 	Theory: Class
		Practical-1 Write the program to Load atleast one of each csv, text and excel file types and display output			Practical Lab
LU2: Index and Filter Data	You will be able to 1. Import data in a python script 2. Index columns using a list of columns 3. Index rows based on a list of index values 4. Index rows based on a conditional statement (mask) 5. Index columns based on a conditional statement (mask) 6. Index columns based on a range of columns 7. Index rows based on a range of index	<ul style="list-style-type: none"> Understanding columns operations and row operations 	Total 30 Hrs Theory: 25 Hrs Practical: 05 Hrs		Theory: Class
		Practical-1 Program to Load all column(s), renames specified column(s), mask specified columns, drop specified column(s) and create column(s) as desired at runtime. Practical-2 Write logic to load even/odd columns, even/odd rows, range of columns, merge columns, select columns and rows based on condition(s) from			Practical Lab

	value	user.			
LU3: Perform Basic Column Level Operations	You will be able to 1. Rename column 2. Apply a function element-wise to a column using “apply” 3. Get value counts of a column 4. Get sum of values in a column 5. Get basic stats of a column (mean/median/standard deviation etc.) 6. Change type of a column 7. Perform a vectorized arithmetic operation on a column 8. Delete a column 9. Duplicate a column 10. Group values of a column and apply an operation on each group	<ul style="list-style-type: none"> Understanding columns operations and row operations. 	Total 20 Hrs Theory: 03 Hrs Practical: 17 Hrs		Theory: Class
		Practical-1 Write a program to get data set information like columns count, rows count, sum of a columns, conditional sum of a column, Detect and change datatypes of columns. Practical-2 Perform basic arithmetic operations over rows and columns, over conditionally selected rows and columns.			Practical Lab
LU4: Handle missing data	You will be able to 1. Count number of missing values in each column 2. Fill missing values with a specific string 3. Fill missing values with mean of the column 4. Delete rows with missing values	<ul style="list-style-type: none"> Understanding columns operations and row operations. 	Total 20 Hrs Theory: 02 Hrs Practical: 18 Hrs		Theory: Class
		Practical-1 Write a program to detect missing values or null fields. Program will delete missing fields. Practical-2 Program will detect datatype of missing value and fill.			Practical Lab

		String value will be replaced with given string and numeric with adjacent rows mean/median or average as asked.			
LU5: Perform string level operations and Regex	You will be able to <ol style="list-style-type: none"> 1. Convert a column to string 2. Divide a column into two based on a separator 3. Check if each row contains a specific substring 4. Extract substring out of each row in a column 5. Check if each row starts with a specific substring 6. Replace a specific substring in each row in a column 7. Change case of a string column 8. Strip spaces from the sides of each row in a column 9. Concatenate a value to each row in a column 10. Concatenate another column with a string 	<ul style="list-style-type: none"> • Understanding of string operations like concatenation, trimming, length calculation and other <hr/> <p>Practical-1 Create a program that can perform following:</p> <ul style="list-style-type: none"> • Select specified column, split selected column from specified delimiter, <p>Practical-2 Create a program to subtract given number of characters from left, right of the column strings. As well as take defined number of columns from starting mentioned index.</p> <p>Practical-3 Create a program to replace specific characters from each row's defined column(s), trim spaces from start and end.</p>	<p>Total 25 Hrs</p> <p>Theory: 20 Hrs</p> <p>Practical: 05 Hrs</p>		<p>Theory: Class</p> <p>Practical Lab</p>

	column elementwise 11. Perform custom operations using “apply”				
LU6: Merge Data	You will be able to 1. Merge two data frames using merge functions 2. Perform different types of joins on two dataframes 3. Concatenate two or more dataframes row wise 4. Concatenate two or more dataframes column wise	<ul style="list-style-type: none"> Understand dataframe formation, loading and reading dataframes 	Total 15 Hrs		Theory: Class
		Practical-1 Code program to use dataframe merge function and use its variants/joins like Left, right, sort and copy. Also write code to concatenate dataframes’ rows and columns in separate functions	Theory: 05 Hrs		Practical Lab
			Practical: 10 Hrs		
LU7: Reshape Data	You will be able to 1. Stack a dataframe 2. Unstack a dataframe 3. Create a pivot table 4. Melt a dataframe 5. Pivot a dataframe	<ul style="list-style-type: none"> Understanding of stacked and unstacked data Understanding of pivot formation 	Total 15 Hrs		Theory: Class
		Practical-1 Write a program to load plane dataframe and apply stack/unstacked function. Practical-2 Write a program to create pivot table with single and multiple lists. Also code dataframe and melt data according to single and multiple variables.	Theory: 03 Hrs		Practical Lab
			Practical: 12 Hrs		

LU8: Apply Row/Cell level Operations	You will be able to 1. Count null values in a row 2. Drop/select specific rows based on a condition 3. Drop/select rows by index 4. Reset index of rows 5. Set a custom index of rows	<ul style="list-style-type: none"> Understanding columns operations and row operations <hr/> Practical-1 Write a program to drop columns and rows against given conditions of null value or given indexes	Total 10 Hrs Theory: 02 Hrs Practical: 08 Hrs		Theory: Class Practical Lab
--	--	--	--	--	------------------------------------

Module: 061900934 Work with Multidimensional Arrays' Manipulation and Computation Package

Objective of the Module: After the completion of this module the candidate will be able to work efficiently with Multidimensional Arrays to perform variety of programming/control jobs.

Duration: 150 hrs. Theory: 30 hrs. Practice: 120 hrs.

Learning Unit	Learning Outcomes	Learning Elements	Duration	Material/Tools Required	Learning Place
LU1: Handle ndarray	You will be able to 1. Read ndarray from pickle file 2. Write ndarray to a pickle file 3. Iterate over arrays 4. Append elements to an ndarray 5. Drop elements from ndarray	<ul style="list-style-type: none"> Describe serializing and de-serializing of objects. Describe pickling in python. Demonstrate importing a pickle file. Describe read and write operation on a pickle file Demonstrate read and write operations on a pickle file. Describe iteration over n-dimensional array. Demonstrate iteration operation over n-dimensional array. Describe append operation on an array. Describe drop operation on an array. Demonstrate append and drop operations, to and from an array. 	<p>Total 20 Hrs</p> <p>Theory: 16 Hrs</p> <p>Practical: 04 Hrs</p>	<ul style="list-style-type: none"> Computer system with python installed Pip package manager Stable version of numpy and pickle Text/Code editing application Terminal/cmd application 	<p>Theory: Class</p> <p>Practical: Lab</p>
		Practical-1 Develop python program to perform following:			

		<ul style="list-style-type: none"> • Import read and write operations on a pickle file. • Iteration operations over n-dimensional array. • Append or extend operations on an array. • Four drop operations from an array 			
LU2: Perform Index ndarray	You will be able to 1. Perform basic slicing and indexing on ndarray 2. Index ndarray using a mask (Boolean array indexing) 3. Index ndarray using integer array indexing	<ul style="list-style-type: none"> • Describe slicing of an array (one-dimensional or two-dimensional). • Describe indexing of an array. • Demonstrate slicing and basic indexing of n-dimensional array. • Describe basic Boolean operations. • Demonstrate Boolean indexing using basic operators. • Describe integer array indexing. • Describe advanced index for column. • Demonstrate selection of arbitrary items based on array dimension. 	<p>Total 20 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 16 Hrs</p>		<p>Theory: Class</p> <p>Practical: Lab</p>
		<p>Practical-1 Develop python program to perform following:</p> <ul style="list-style-type: none"> • Basic slicing and indexing of n-dimensional 			

		<p>array using ndarray</p> <ul style="list-style-type: none"> • Boolean indexing using basic operators • Boolean indexing using advance operations • Selection of arbitrary items based on array dimension. 			
<p>LU3: Operate on ndarray</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Perform binary operations on arrays 2. Perform string operations on arrays 3. Perform comparison operations on arrays 4. Change type of an array 5. Split arrays (split, dsplit, vsplit, hsplit) 6. Tile arrays 7. Rearrange array (reshape, roll, flip) 	<ul style="list-style-type: none"> • Describe binary operation on arrays • Demonstrate bitwise binary operation on arrays • Describe string operation on arrays • Demonstrate various string operation on arrays • Demonstrate various comparison of arrays. • Describe various array types. • Demonstrate changing the type of an array. • Describe various split operations that can be performed on ndarrays • Demonstrate various split operations that can be performed on ndarrays • Describe basic functionality of tile arrays • Demonstrate how to construct tile 	<p>Total 25 Hrs</p> <p>Theory: 05 Hrs</p> <p>Practical: 20 Hrs</p>	<p>Theory: Class</p> <p>Practical: Lab</p>	

		<p>array</p> <ul style="list-style-type: none"> Describe operations for rearranging arrays Demonstrate various functions to rearrange array 			
		<p>Practical-1</p> <p>Develop a python program to perform following:</p> <ul style="list-style-type: none"> Bitwise binary operation on arrays Various string operation on arrays Comparison of arrays. Change the type of an array. <p>Practical-2</p> <p>Develop a python program to perform following:</p> <ul style="list-style-type: none"> Split operations that can be performed on ndarrays Construct tile array Rearrange array 			
<p>LU4: Reshape ndarray</p>	<p>You will be able to</p> <ol style="list-style-type: none"> Change dimensions with “reshape” Flatten array with “ravel” Move axis of an array Roll axis of an array Swap axes of an array 	<ul style="list-style-type: none"> Describe basics of reshape method using numpy Describe basics of ravel method using numpy Demonstrate reshape and ravel operation on ndarrays Describe basics of move axis operation of an array to a new position Describe the 	<p>Total 25 Hrs</p> <p>Theory: 05 Hrs</p> <p>Practical: 20 Hrs</p>	<p>Theory: Class</p> <p>Practical: Lab</p>	

	<p>6. Take transpose of an array</p> <p>7. Broadcast an array</p>	<p>basics of roll axis operation</p> <ul style="list-style-type: none"> Describe the basics of swap axis operation Demonstrate move, roll, swap axis operations on arrays Describe the purpose of transpose operation Demonstrate transpose operation on arrays Describe the broadcast operations on an array Demonstrate Broadcasting on ndarrays 			
		<p>Practical-1</p> <p>Develop a python program to:</p> <ul style="list-style-type: none"> Perform reshape and ravel operations on ndarrays Move, roll, swap axis operations on arrays Transpose operation on arrays Broadcasting on ndarrays 			
<p>LU5: Merge ndArrays</p>	<p>You will be able to</p> <ol style="list-style-type: none"> Concatenate arrays Stack arrays Stack 1D arrays as columns in a 2D array (column stack) Perform stacking 	<ul style="list-style-type: none"> Describe concatenate functionality of arrays in numpy package Demonstrate concatenation of arrays using numpy package Describe various stacking operations for 	<p>Total 20 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 16 Hrs</p>		<p>Theory: Class</p> <p>Practical: Lab</p>

	on particular axes (dstack, hstack, vstack)	<p>arrays</p> <ul style="list-style-type: none"> • Demonstrate staking of arrays using numpy • Describe column stacking • Demonstrate stacking 1D array as columns in 2D array • Demonstrate stacking on various axes 			
		<p>Practical-1</p> <p>Develop a python program that can:</p> <ul style="list-style-type: none"> • Concatenates multiple arrays using numpy package • Perform stacking of arrays • Perform column stacking • Perform stacking on various axes 			
LU6: Process Text Data	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Read text documents into variables 2. Tokenize text documents 3. Count number of unique words in a document 4. Convert a text document into a label encoded array 5. Encode a document phrase using one hot encoding 	<ul style="list-style-type: none"> • Describe basic list of packages available for processing text in python • Demonstrate installing relevant text processing package • Describe text tokenization concept • Demonstrate reading text document in python and perform tokenization • Describe basic counting principles • Demonstrate counting number of unique words in documents 	<p>Total 20 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 16 Hrs</p>		<p>Theory: Class</p> <p>Practical: Lab</p>

		<p>using relevant python package</p> <ul style="list-style-type: none"> • Describe label encoding • Describe on hot encoding • Describe packages available for label 			
		<p>Practical-1 Develop a python program that can:</p> <ul style="list-style-type: none"> • Reads a text document and performs tokenization • Count the number of unique words in a text document <p>Practical-2 Develop a python program to:</p> <ul style="list-style-type: none"> • Convert text document to label • Perform one hot encoding on text data 			
<p>LU7: Handle new sources of Data</p>	<p>You will be able to</p> <ol style="list-style-type: none"> 1. Read Audio data as numpy array 2. Read Image data as numpy array 3. Read LIDAR data as numpy array 4. Read Time Series data as numpy array 	<ul style="list-style-type: none"> • Install PyAudio package • Describe basic functionality of PyAudio • Demonstrate recording audio with PyAudio • Demonstrate conversion of audio data buffer to numpy ndarray • Install python package for image processing • Describe basic functionality of image processing 	<p>Total 20 Hrs</p> <p>Theory: 04 Hrs</p> <p>Practical: 16 Hrs</p>		<p>Theory: Class</p> <p>Practical: Lab</p>

		<p>package</p> <ul style="list-style-type: none"> • Demonstrate loading image data • Demonstrate conversion of image data to numpy ndarray • Install suitable python package to process LIDAR data • Demonstrate loading LIDAR data • Demonstrate conversion of LIDAR data to numpy ndarray • Install suitable python package to process time series data • Demonstrate loading time series data • Demonstrate conversion of time series data to numpy ndarray 			
		<p>Practical-1 Develop a python program to:</p> <ul style="list-style-type: none"> • Load audio data and convert it to numpy ndarray • Load image data and convert it to numpy ndarray • Load LIDAR data and convert it to numpy ndarray • Load time series data and convert it to numpy ndarray 			

General assessment guidance for *Artificial Intelligence Data Technician*

Good practice in Pakistan makes, use of sessional and final assessments, the basis of which is described below. Good practice by vocational training providers in Pakistan, is to use a combination of these sessional and final assessments, combined to produce the final qualification result.

Sessional assessment is going on all the time. Its purpose is to provide feedback on what students are learning:

- To the student: to identify achievement and areas for further work
- To the teacher: to evaluate the effectiveness of teaching to date, and to focus future plans.

Assessors need to devise sessional assessments for both theoretical and practical work. Guidance is provided in the assessment strategy

Final assessment is the assessment, usually on completion of a course or Level, which says whether or not the student has "passed". It is – or should be – undertaken with reference to all the objectives or outcomes of the course, and is usually fairly formal. Considerations of security – ensuring that the student who gets the credit is the person who did the work – assume considerable importance in final assessment.

Methods of assessment

For lessons with a high quantity of theory, written or oral tests related to learning outcomes and/ or learning content can be conducted. For workplace lessons, assessment can focus on the quality of planning the related process, the quality of executing the process, the quality of the product and/or evaluation of the process.

Methods include direct assessment, which is the most desirable form of assessment. For this method, evidence is obtained by direct observation of the student's performance.

Examples for direct assessment of Artificial Intelligence Data Technician Lev-1-4 include:

- Demonstrations, for example demonstrating recording audio with PyAudio and conversion of LIDAR data to numpy ndArra.
- Paper-based tests, such as multiple choice or short answer questions on data processing and program installation.

Indirect assessment is the method used where the performance could not be watched and evidence is gained indirectly. Indirect assessment should only be a second choice. (In some cases, it may not even be guaranteed that the work products were produced by the person being assessed.)

Examples for direct assessment of Artificial Intelligence Data Technician Lev-1-4 include:

- Portfolio, for example student is asked to bring his previous scripts of python written by him.

Principles of assessment

All assessments must meet all the following principles, regardless of the method of assessment used to evidence learners' attainment.

All assessments must produce outcomes that are:

1. Valid: the assessment evidence meets all assessment criteria and all learning outcomes
2. Authentic: all the work is the learner's own
3. Reliable: assessment evidence is consistent and generates outcomes that would be replicated were the assessment repeated
4. Current: assessment evidence is up-to-date
5. Sufficient: enough work is available to justify the credit value, and to enable a consistent and reliable judgement about the learner's achievement
6. Comparable: all assessment evidence is comparable in standard between assessments within a unit/qualification, and between learners of the same level
7. Manageable: all assessment places reasonable demands on all learners
8. Fair and minimize bias: assessments are fair to all learners irrespective of their characteristics (for example, age, gender etc.)

Assessment strategy for Artificial Intelligence Data Technician Level 4 Curriculum

This curriculum consists of 10 modules:

Module-01	Contribute to Work Related Health and Safety (WHS) Initiatives
Module-02	Comply with Workplace Policy and Procedures
Module-03	Perform Advanced Communication
Module-04	Manage Human Resource Services
Module-05	Scrape data from the web
Module-06	Process Images through Image Processing software
Module-07	Work with Data Manipulation Toolkit
Module-08	Work with Multidimensional Arrays' Manipulation and Computation Package
Module-09	Develop Advance Computer Application Skills
Module-10	Develop Entrepreneurial Skills

Sessional or Developmental assessment

The sessional/developmental assessment shall be conducted after completion of each module in two parts: theoretical assessment and practical assessment.

Theoretical assessment for all learning modules must consist of a written paper lasting at least 30 minutes per module. This can be a combination of multiple choice and short answer questions.

For practical assessment, all procedures and methods for the modules must be assessed on a sessional basis. Guidance is provided below under Planning for assessment.

Final assessment

Final assessment shall also be in two parts: theoretical assessment and practical assessment.

For the final practical assessment, each student shall be assessed over a period of 4-5 hours' session. During this period, each student must be assessed on his ability to perform a complete job for all Technical and functional modules.

Generic modules shall be assessed comprising with other modules at the time of final assessment. Practical work for this module could be assessed on a sessional basis.

Planning of assessment.

Planning of assessment will plan by the assessment Centre as per CBT/A policy. But for development assessment it could be plan by the Trainer during the course.

As for final assessment as concern, certified assessor must be contacted and the assessor must meet the needs of the students and the training provider. For example, where two assessors are conducting the assessment, there must be a maximum of five students per assessor. In this example, a group of 20 students shall therefore require assessments to be carried out over a four-day period. For a group of only 10 students, assessments would be carried out over a two-day period only or it could be formulated as per CBT/A Centre policies.

Complete list of tools and equipment – Non Consumable

S. No	Description	Quantity
1	Printer	5
2	High performance Computer system with <ul style="list-style-type: none"> • GPU(s) • Audio card • High speed high capacity storage • Drivers • Compatibility with Python, OpenCV 	20
3	Software and Libraries packages: <ul style="list-style-type: none"> • Microsoft Office • Python Software Package • BeautifulSoup Python Library • Request Python Library • OpenCV software package (latest version) • Pip package manager • Stable version of pandas • Stable version of numpy and pickle 	20
4	High quality sound system	20
5	High resolution display	20
6	Software application(s) for image manipulation	20
7	Software application(s) for audio editing	20
8	Software application(s) for video manipulation and processing	20
9	Noise absorbers	20
1	Scanner	05
1	Digital Camera	05
1	Internet facility	-
1	Virtual environment package	20

Complete list of tools and equipment - Consumable

S. No.	Items
1.	Different Tags and Locks
2.	Paper rim
3.	Process SOPs
4.	Equipment Maintenance Manuals
5.	Log Book
6.	Handbooks
7.	Design Books/ Sheets
8.	Pencils
9.	Erasers
10.	Pencil Sharpeners
11.	Paper Cutter
12.	Scissors
13.	Color Pencils
14.	White chart paper
15.	Brown Sheets
16.	White Board Markers (red, blue, green, black)
17.	Permanent markers (black)
18.	File covers

Credit values

The credit value of the National Certificate Level 4 in Artificial Intelligence Data Technician is defined by estimating the amount of time/ instruction hours required to complete each competency unit and competency standard. The NVQF uses a standard credit value of 1 credit = 10 hours of learning (Following TVET guidelines).

The credit values are as follows:

Code	Name of Duty or (Module)	Level	Credit	Category
	Contribute to Work Related Health and Safety (WHS) Initiatives	4		Generic
	Comply with Workplace Policy and Procedures	4		Generic
	Perform Advanced Communication	4		Generic
	Manage Human Resource Services	4		Generic
	Scrape data from the web	4	13	Technical
	Process Images through Image Processing software	4	12	Technical
	Work with Data Manipulation Toolkit	4	15	Technical
	Work with Multidimensional Arrays' Manipulation and Computation Package	4	15	Technical
	Develop Advance Computer Application Skills	4		Generic
	Develop Entrepreneurial Skills	4		Generic

