AI ML Using Python , C# and Azure AI

Syllabus

www.questpond.com

AI ML Fundamentals

AI ML Fundamentals	
What is AI and ML?	
How to Human train:- Features and Labels. Alphabet Image	
Model :- Connection between Features and Labels. Human thinking Mapping.	
Defining Features, Labels and Models ?.	
Putting things together the Pipeline. FLAP M (Features , Labels , Algo , Pipeline , model)	
Format of Data for Training machines Data Format	
Regression:- Finding relationship between Numerical data.	
Understanding Regression Analysis with a simple data using Excel.	
Introduction to ML.NET C# , understanding MLContext , MLData and MKL components.	
OLS Ordinary Least Squares and SDCA Stochastic Dual Coordinate Ascent.	
Single Prediction vs Test Data	
R Square and RMSE (Root Mean Squared)	
Checking which is best Algo ?	
OLS with polynomial data	
AutoML	
Performance comparison of scratch and Online training models.	
Feature Engineering.	
Linear , Non-Linear and Seasonal.	
Supervised Learning and Unsupervised Learning.	
Clustering Algorithms and KMeans	
What is MLP ?	
Vector , Tokens , Encoding,Embedding , Transformer , BERT and GPT ?	
MLP Encodings :- One-Hot Encoding , BOW , TF IDF , Word and Transformer Embeddings	
Cosine and Euclidean	
Lab 1 :- Simple Example of Linear Regression using ML.NET using single prediction.	
Lab 2 :- Simple Example of Linear Regression using ML.NET using test data.	
Lab 3 :- Simple Example of Linear Regression checking R Square and RMSE score	
Lab 4 :- Running Example with Linear OLS and SDCA.	
Lab 5 :- AUTOML	
Lab 6 :- Load Huge File and check AUTOML Suggestions and check Accuracy	
Lab 7 :- Saving model and retraining through Live training (SDCA and Online Gradient Descent)	
Lab 8 :- Binary / Logistic regression.	
Lab 9 :- Multi Class Classification	
Lab 10 :- Simple Clustering Example using KMeans.	
Lab 11 :- Understanding One Hot Encoding.	
Lab 12 :- Simple Example of BOW	
Lab 13 :- Simple example of TF-IDF	

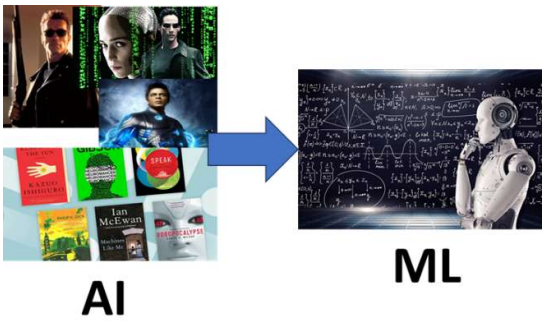
Lab 14 :- Example of WordEmbedding using GloVe50D and similarity checking using COSINE and Euclidean	
Lab 15 :- Simple BERT Example of Interview Question lookup.	
Lab 16 :- GPT Example not working.	
Lab 17 :- ChatGPT using Prompt fundamentals.	
Unsupervised: - Clustering and Dimensionality reduction	
Q Learning and Deep Q networks.	
LLM's and LLM's :- GPT, Claude , LLAMA ,Mistral	
Setting Context :- MCP Model Context Protocol	
Agent and Generative AI	
Orchestration using N8N / ?.	
Python Basics	
Getting your first Python Program Running.	
Python Data types, strings, statements and functions.	
Object Oriented Programming and Error handling	
Modules and Packages	
Numpy and Pandas	
SciKit , Tensor flow , Pytorch, SciPy , NLTK , OpenCV	
Prompt Engineering	
Prompt Engineering (Personal , Task , Context , Constraints and Format)	
JSONL files :- Persona , System , Rules and Settings.	
Azure AI	
Identifying Machine Learning Types	
AutoML :- No Code machine learning.	
Computer vision work loads.	
NLP workloads.	
Generative AI work loads	
Azure Open AI service	
Project	
Nifty Prediction :- Simple Seasonal Prediction Lag way of prediction Predicting Wars , Black swan What does it take to create a perfect prediction ?	

Contents

What is AI ML ?

AI (Artificial Intelligence) is the broader field that makes machines think and act like humans.

ML (Machine Learning) implements AI by teaching machines to learn from data and improve automatically. WITH OUT PROGRAMMING CHANGES



Explain Features, Labels and models?

Features are input, Label target output and Model mathematical mapping from Features to Label.

What is Pipeline?

It is series of steps which needs to be executed to get the final model.

What are MKL components?

MKL components are optimized math libraries (Intel® Math Kernel Library) that accelerate numerical computations. If this is not there it will use .NET Math libraries.

Explain Regression ?

Regression is a machine learning technique that predicts a continuous numeric value based on input data.

What is OLS and SDCA?

OLS finds best line in ONE Shot. Imagine you have dots on your graph, OLS tries to draw one line which fits all DOTS with least total error.

SDCA starts with a random guess, checks , adjusts and repeats many times until its gets the best line.

Explain R SQUARE and RMSE ?

R^2 :- 0 → 1 (higher better) ,Fit quality (pattern correctness)

RMSE :- 0 → ∞ (lower better),Error size (prediction closeness)

Different types of Regression Algorithms and When to use them?

Algorithm	***Type***	***When to Use***	***Incremental***
SDCA Regression	Linear	Large datasets, fast training, general-purpose regression	Yes
SDCA Logistic Regression	Linear (classification)	When target is 0/1 but treated in regression-style pipelines	Yes
Online Gradient Descent (OGD)	Linear	Streaming data, frequent retraining, simple linear relationships	Yes

Algorithm	***Type***	***When to Use***	***Incremental***
FastTree Regression	Gradient boosting	Nonlinear patterns, high accuracy, tabular data	No
FastForest Regression	Random forest	Robust to noise, nonlinear relationships, easy tuning	No
LightGBM Regression	Gradient boosting (GOSS)	Best accuracy on complex datasets, many features	No
Gam Regression	Generalized additive model	Interpretable, captures smooth nonlinear trends	No
Poisson Regression	Linear (count data)	Prediction of counts (events per interval)	No
Tweedie Regression	Linear (compound Poisson–Gamma)	Insurance, rainfall, zero-inflated continuous values	No
Linear Regression (OLS)	Linear	Small datasets, simple relationships, interpretable	No

Supervised Learning vs Unsupervised Learning?

Supervised learning:- Model is trained using labelled data (input + correct output), so it learns to predict the right answer. Linear Regression , Logistic regression , Multi class
Unsupervised learning:- Model is trained using only input data (no labels), and the algo discovers patterns. Kmeans , Clustering

Logistic (Binary) and Multi-class Regression?

Logistic regression predicts Yes and No as outcome.
Multi-class Regression predicts category.

Explain Clustering ?

Clustering groups similar data points together on their patterns without labels.

What is KMeans ?

It's a clustering algorithm which divides data into K groups by placing each point into the nearest cluster center.

What is NLP?

NLP (Natural Language Processing) is the field of AI that enables computers to understand human language.

How Does NLP Work?

Its converts (Encodes) text to numbers.

What are vectors , token, Encoding , Tokenization and Embedding?

Vectors are single or group of numbers which represents a text.

Shiv is a Trainer – Vector → [1,0,0,1]

Other word of Vectorization is Embedding / Encoding.

Token represents a piece of word from text. Tokenization is the process breaking text into small units called tokens so that model can process it.

I love AI and facebook – Tokenization → "I" "LOVE" "AI" "AND" "FACE" "BOOK"

Explain Transformer ?

A Transformer is type of neural network which understand relationships between words in a sentence AT ONCE and uses ATTENTION technique to get context.

"I went to the **bank ATM** and later relaxed near **river bank**."

"When I was **eating Apple** , **apple** released a new **IPhone**".

What is BERT and GPT ?

BERT (Bidirectional Encoder Representations from Transformers) :- Understands text by reading both left and right context — best for understanding meaning.

GPT (Generative Pre-trained Transformer) :- Predicts next words from left-to-right — best for generating sentences.

Explain One-Hot Encoding , BOW , TF IDF , Word and Transformer Embeddings?

One-Hot Encoding :- Every word becomes a unique 0/1 vector.
Bag-of-Words (BoW) :- Counts how often each word appears.
TF-IDF :- Highlights important words and downplays common ones.
Word Embeddings :- Words with similar meaning have similar numbers.
Transformer Embeddings :- Meaning depends on full sentence context . (Attention is all you need)

How is GPT connected with ChatGPT ?

What is Prompt Engineering ?

Giving Clear instruction to the AI model to get the best possible output.

What are Roles

Role :- Tell AI Who to be ?

System :- Gives instruction how AI should behave.

User :- The Human input.

Assistant :- AI's reply.

https://platform.openai.com/chat

What is RAG ?

Indexing -- Retrieve – Augment – Generate

Document D1	*The child makes the dog happy* the: 2, dog: 1, makes: 1, child: 1, happy: 1
Document D2	*The dog makes the child happy* the: 2, child: 1, makes: 1, dog: 1, happy: 1

↓

	child	dog	happy	makes	the	BoW Vector representations
D1	1	1	1	1	2	[1,1,1,1,2]
D2	1	1	1	1	2	[1,1,1,1,2]

Sentence 1 : The car is driven on the road.

Sentence 2: The truck is driven on the highway.

TF-IDF Recap

TF-IDF is a way to measure how important a word is in a document relative to a collection of documents (corpus). It has two parts:

1. **TF (Term Frequency)** – How often a word appears in a document.

$$TF(t,d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}$$

2. **IDF (Inverse Document Frequency)** – How rare or common a word is across all documents.
   The rarer a word, the higher its IDF. It is defined as:

$$IDF(t) = \log \frac{N}{1 + DF(t)}$$

   where:

   - N = total number of documents
   - $DF(t)$ = number of documents containing term t
   - $1+$ is added to avoid division by zero

Why logarithm?

Word	TF		IDF	TF*IDF	
	A	B		A	B
The	1/7	1/7	log(2/2) = 0	0	0
Car	1/7	0	log(2/1) = 0.3	0.043	0
Truck	0	1/7	log(2/1) = 0.3	0	0.043
Is	1/7	1/7	log(2/2) = 0	0	0
Driven	1/7	1/7	log(2/2) = 0	0	0
On	1/7	1/7	log(2/2) = 0	0	0
The	1/7	1/7	log(2/2) = 0	0	0
Road	1/7	0	log(2/1) = 0.3	0.043	0
Highway	0	1/7	log(2/1) = 0.3	0	0.043

1. TF — Term Frequency (How often does a word appear?)

If the word “cat” appears 2 times in a sentence of 10 words:

TF = 2 / 10 = 0.2

2. IDF — Inverse Document Frequency (How rare is this word?)

If a word appears in **every** sentence, it is not special → low IDF
If it appears in **one** sentence only → high IDF

IDF = log(total_documents / docs_containing_word)

TF–IDF = TF × IDF

So TF–IDF gives:

- **High score** → word is important for that sentence
- **Low score** → word is common or not useful

- **TF** tells what’s important *inside* a sentence
- **IDF** tells what’s important *across all sentences*
- **TF-IDF** combines both → so rare meaningful words get more weight

Alphabet Image

Features and Labels

Feature	Label
Age	Insurance
10	2000
20	4000
30	6000
40	8000
50	10000
60	12000
70	14000
80	16000
90	?

Human thinking mapping with Maths

Human Thinking Style	Human Example	Computer / Mathematical Model
Linear Thinking	*"If I study more hours, I will get more marks."*	**Linear Regression** → fits a straight line $y = mx + c$
Non-Linear Thinking	*"Weight loss is fast initially, then slows down."*	**Polynomial / Exponential Regression** → curved function fitting
Threshold-Based Decision	*"If fever > 100°F, then take medicine, else don't."*	**Decision Trees** → rules based on splitting conditions
Similarity-Based Reasoning	*"This handwriting looks like Rahul's handwriting."*	**k-Nearest Neighbors (KNN)** → compares **distance** between examples
Weighing Multiple Clues	*"When buying a phone, battery matters more than color."*	**Logistic/Linear Regression Weights** → features get **importance scores**
Learning from Experience / Feedback	*"I adjust my cricket shot based on where the ball landed last time."*	**Reinforcement Learning** → learns by maximizing **reward** over time
Seeing Patterns as a Whole	*"I recognize my friend instantly from far away."*	**Neural Networks / Deep Learning** → layered pattern recognition matrices
Grouping / Categorization	*"These 3 items look simil… they must belong together."*	**Clustering (K-Means, DBSCAN)** → groups by minimizing internal distances

Polynomial code for OLS

```
var pipeline = mlContext.Transforms.CustomMapping<InsuranceData, PolynomialData>(
 (input, output) => { output.Age = input.Age; output.Age2 = input.Age * input.Age; output.Premium = input.Premium; },
 contractName: "PolynomialMapping")
 .Append(mlContext.Transforms.Concatenate("Features", "Age", "Age2"))
 .Append(mlContext.Regression.Trainers.Ols(labelColumnName: "Premium", featureColumnName: "Features"));
```

Nifty model prediction

Method Type	**Difficulty**	**Accuracy**
Lag-based ML models	Easy	Good
Technical indicators	Easy	Medium

Method Type	Difficulty	Accuracy
Macro models (Gold, USD/INR, Crude, VIX)	Medium	Good
ARIMA / Holt / Prophet	Easy	Medium
LSTM / GRU / Deep Learning	Medium–Hard	Excellent
Hybrid models (ML + Macro + Indicators)	Hard	Very high
Sentiment + news models	Medium	Good in events
Trend-following rules (EMA, SuperTrend)	Very easy	Medium
Regime detection (HMM, GARCH)	Medium	Good risk control
Ensemble models (combine many models)	Medium	Excellent

Question :- Why the same data, same formula, giving us 2 different values? excel and console?

Answer :- ML.NET we used OLS , Excel also uses OLS but they are different engines. ML.NET uses MKL (Intel) while excel uses internal maths libraries of Microsoft. Excel floating point precision and C# is different so differences arise from there as well.

Question :- What is R mean?

Answer :- R Square check if the algorithm fits correctly. RMSE says how much correct is the value (prediction closeness)

Question :- So the score attribute here identify the best match and share the results?

Answer :- For regression the score is the predicted output but if its classification its confidence value.

Question :- can we define all feature in one concatenate statement..?

Answer :- Yes you can.

Question :- Which VS version is best for ML/AI ?

Answer :- VS 2022 should be fine. Some students said VS code absolutely fine.

Question :- Is visual studio community ok or visual studio code?

Answer :- In training i am using VS community. VS code is also fine.

Question :- Why ML.NET as we have dotnet logics to achieve?

Answer :- If we do not use ML.NET then we have write all these from scratch. Its huge task just think.

Question :- Insurance limit & Existing/New customer will be Feature or Label ?

Answer :- Anything which is needed to train the algorithm will be a feature. For Insurance limit generally thinking i will say its a label.

Question :- How to identify features and labels in the data ?

Answer :- Labels are mostly what you want to predict (output) and to get the output what are the inputs (features) we need. So if you know your input and output you should be able to identify features and labels.

Question :- Is this ML.NET code a bit complicated?

Answer :- Yes first it would look like, try to follow the video and look at fluent API code. If you understand Fluent API code you should do well.

Question :- It seems like there are known trade offs for all algorithm, like speed, accuracy etc. Good to know some globally accepted set of rules before designing pipeline

Answer :- Yes there are tradeoffs in terms of accuracy , speed. Its a iterative process.

Question :- The maths part will be done by the code .and algorithm will be interpretd by the code

Answer :- Maths part is called by the Algorithm and we just use the model.

Question :- From your experience in real world cases, would you use multiple models and perhaps get the mean (or compare) as a way to guarantee accuracy (as much as possible) (in addition to R Sq). And do you retrain as you get more data through actual use?

Answer :- Getting Accuracy is an ART , Trial and error , out of the box thinking. I would not try mean as that would lead to cocktail but i will try weird custom code ways of getting proper output.

Question :- so algorithm is generic and model is derived from data and algorithm

Answer :- Yes perfectly thought , Algorithm is generic but when filled with values it makes sense.

Y = MX+C is Linear Regression Algorithm , Y = 0.5*X + 100 is a model.

Question :- Hope i can consider Insurance as Feature and Age as Label as well?

Answer :- Yes you can but think will some one predict age.

Question :- Sir, can you explain linear, polynomial & exponential concepts

Answer :- Linear relationships change at a constant rate, forming a straight line; polynomial relationships change at varying rates, creating curves that can rise and fall; and exponential relationships grow or decay rapidly, multiplying by a constant factor over equal intervals.

We will discuss linear, polynomial & exponential difference in class also.

Question :- When to use ML.NET over Azure AI ?

Answer :- ML.NET is only when you are doing something Custom , small and simple. For Enterprise level ML you will need ready made AI tools like Azure AI , Open AI.

Question :- Could you please suggest good book for ML.Dotnet ?

Answer :- You can check out Programming ML.NET by Dino Esposito and Francesco Esposito. Problem with books is framework changes and books are updated late. Better would be live classes as you get the recent version and code.

Question :- This are all AI not GEN AI right ?

Question :- AI vs Gen AI?

Answer :- Answering both questions in ONE go. This is not GEN AI its basic AI. Gen AI creates new content, like text or images, based on patterns in data.

Question :- Do we need to know all the algorithms? I am what each algorithm does.

Answer :- A Basic GIST should be ok , do not get in too much Maths and understanding.

Question :- what should be weekly plan . should we practice this code ?

Answer :- Just practice the code read , discuss on our discord group , WhatsApp channels

Question :- when this recording will available ?

Its already done.

Question :- Will we get the recording of the session to follow the steps and try it out ?

Yes you will in teachables and if you are in Cohort then in podia.