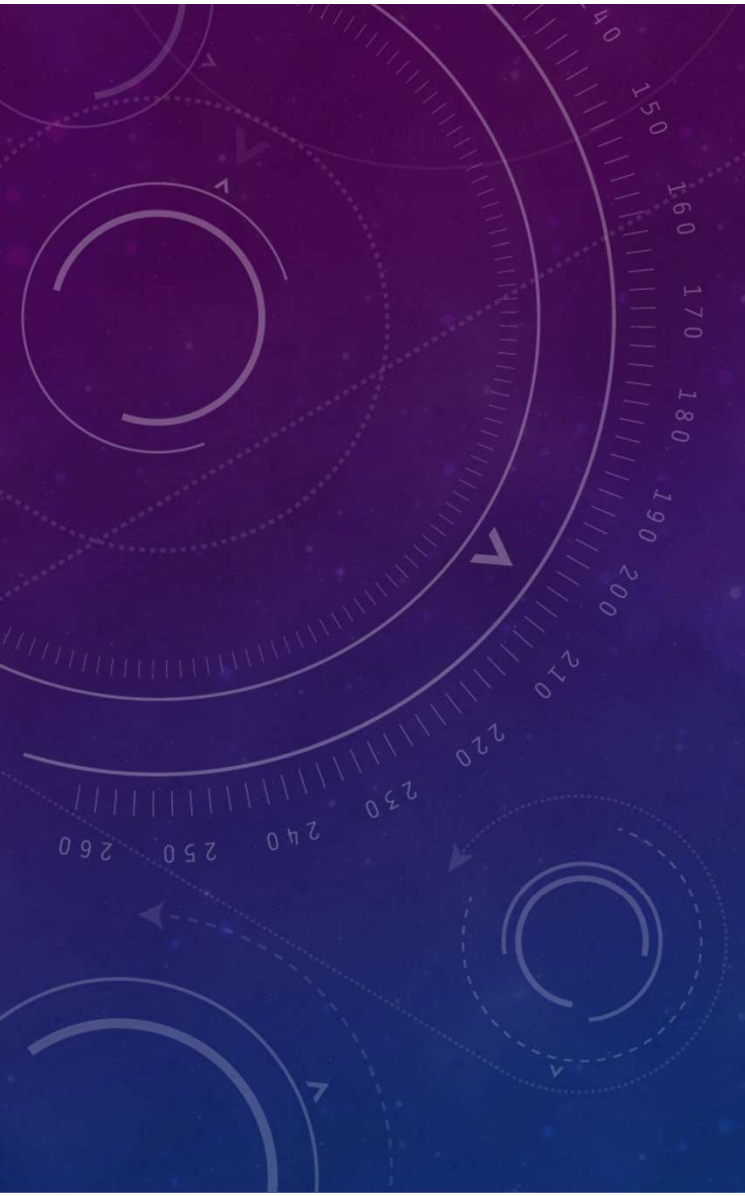


MACHINE LEARNING WITH .NET

Asmaliza Ahzan @ Emma

asmaliza@iverson.com.my

INTRODUCTION

- Name
- Job Title/Role/Descriptions
- Experiences in programming, ML/AI, data analysis or anything related
- Expectations from this course

LOGISTICS

- 3 days course
- Session: 9am-5pm
- 10-15 minutes breaks
- 1 hour lunch break
- Important: Make sure to sign your attendance every day.

GETTING STARTED WITH MACHINE LEARNING AND ML.NET

CHAPTER 1

MACHINE LEARNING

- ML is the study of computer algorithms that can improve automatically through experience and by the use of data.
- It is part of artificial intelligence (AI).
- ML algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so.
- ML algorithms are used in a wide variety of applications, such as medicine, email filtering, speech recognition, and computer vision.

TYPES OF LEARNING

- Supervised learning
  - Supervised machine learning can take what it has learned in the past and apply that to new data using labelled examples to predict future patterns and events. It learns by explicit example.
- Unsupervised learning
  - Supervised learning tasks find patterns where we have a dataset of “right answers” to learn from. Unsupervised learning tasks find patterns where we don’t. This may be because the “right answers” are unobservable, or infeasible to obtain, or maybe for a given problem, there isn’t even a “right answer” per se.
  - Unsupervised learning is used against data without any historical labels.

ML ALGORITHMS

- Binary Classification
- Regression
- Anomaly Detection
- Clustering
- Matrix Factoring

BINARY CLASSIFICATION

- Supervised learning algorithm.
- Task involves classifying the elements of a set into two groups (true/false) on the basic of a classification rule.
- Typical scenarios;
  - Medical testing to determine if a patient has certain disease or not.
  - Quality control in industry, deciding whether a specification has been met.
  - In information retrieval, deciding whether a page should be in the result set of a search or not.

REGRESSION

- Supervised learning algorithm.
- A set of statistical processes for estimating the relationships between a dependent variable (outcome/label) and one or more independent variables (predictors/features).
- Regression analysis is widely used for prediction and forecasting.
- Example scenarios;
  - Weather forecasting
  - Sales performance analysis
  - Stock market prediction
  - House pricing

ANOMALY DETECTION

- The identification of unexpected or rare items, events or observations which raise suspicions in data.
- Typically the anomalous items will translate to some kind of problem such as bank fraud, a structural defect, medical problems or errors in text.
- Anomalies (spiked/change points) also referred to as outliers, noise, deviations and exceptions.

CLUSTERING

- Unsupervised learning algorithm.
- An algorithm that look for patterns in data, such as groups of customers based on their behaviour.
- During training, data is grouped based on the features, and then during the prediction, the closest match is chosen.
- Example;
  - Sorting music files.
  - Predicting customer preferences.

MATRIX FACTORIZATION

- An algorithm to provide recommendation.
- This algorithm is tailored to problems where historical data is available and the problem to solve is predicting a selection from that data.
- Example: Netflix recommendations.

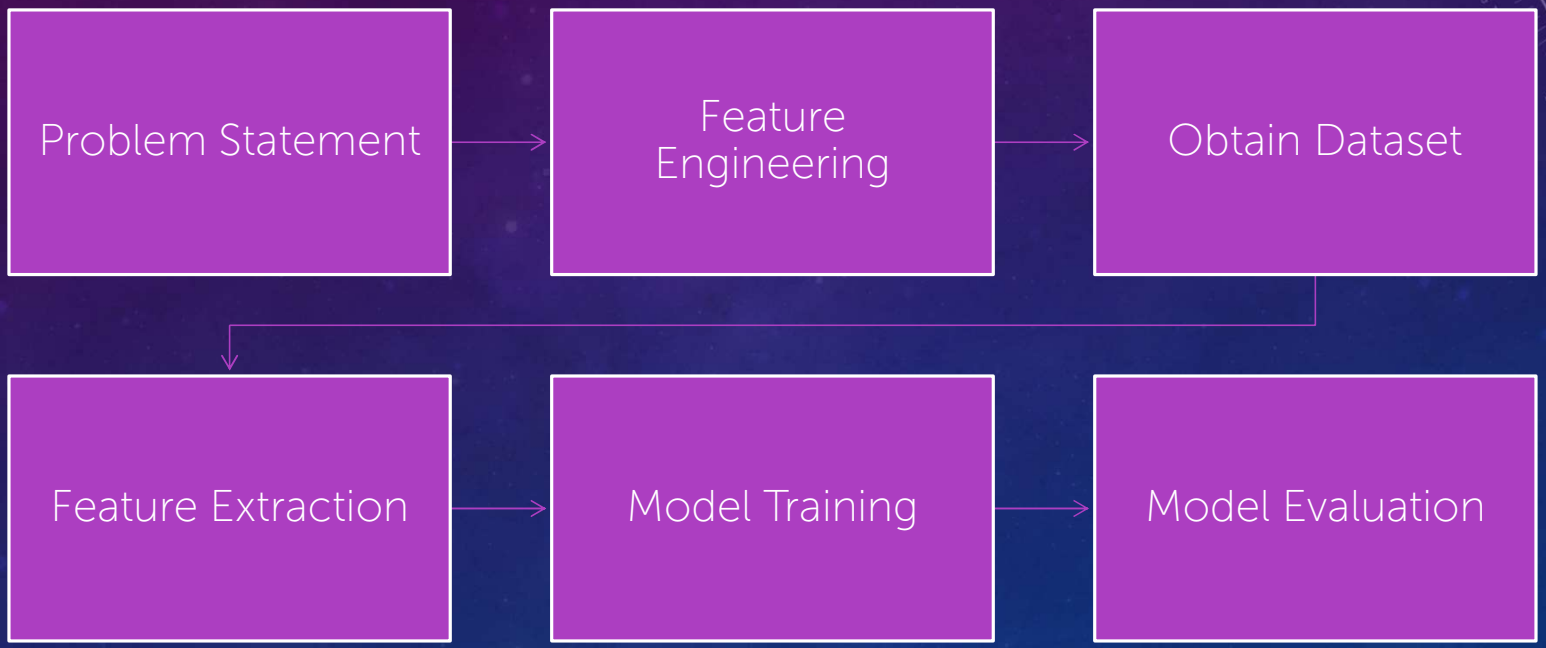
WHAT IS ML.NET AND HOW DOES IT WORK?

- Microsoft ML framework.
- First announced and released in May 2018.
- ML.NET gives the ability to add ML to .NET applications (offline/online).
- Allows automatic predictions using the data available to your application.
- Open-sourced with MIT licensed on GitHub.
- Targeted to C# and F# developers.

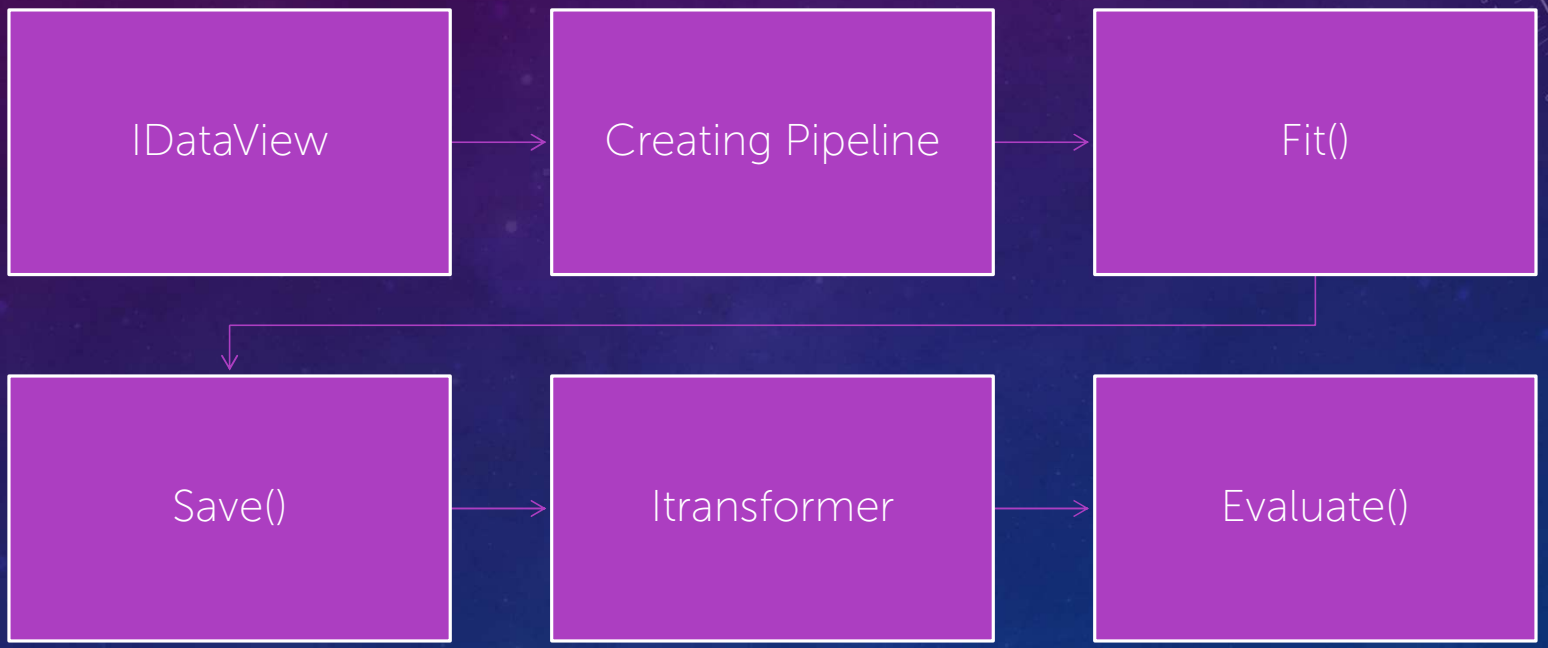
MACHINE LEARNING MODEL

- Central to ML.NET is a ML model.
- The model specifies the steps needed to transform input data into a prediction.
- With ML.NET, a model can be trained by specifying an algorithm.
- Once model is prepared, it can be added to application to make the predictions.

THE MODEL BUILDING PROCESS



HIGH-LEVEL ARCHITECTURE OF ML.NET



COMPONENTS OF ML.NET

- An ML.NET application starts with an MLContext object.
- This singleton object contains catalogs.
- A catalog is a factory for data loading and saving, transforms, trainers, and model operation components.
- Each catalog object has methods to create the different types of components.

ML.NET ARCHITECTURE

Category	Catalogs
Data loading and saving	DataOperationsCatalog
Data preparation	TransformCatalog
Training algorithms	BinaryClassificationCatalog, MulticlassClassificationCatalog, AnomalyDetectionCatalog, ClusteringCatalog, ForecastingCatalog, RankingCatalog, RegressionCatalog, RecommendationCatalog, TimeSeriesCatalog
Model usage	ModelOperationsCatalog

THE END