**TASK 1**

1. **HOW DATA SCIENCE IS REQUIRED FOR MACHINE LEARNING?**

Machine learning and statistics are parts of data science. So there’s plenty of relations between data science and machine learning.The machine learning algorithms train on data delivered by data science to become smarter and more informed in giving back business predictions.Thus, ML algorithms depend on the data; they won't learn without using it as a training set.But sure, data science applies to much more than machine learning. In DS, information may or may not come from a machine or mechanical process. Survey data, for example, can be collected manually.Sometimes it may have nothing to do with learning.The main difference lies in the fact that data science covers the whole spectrum of data processing. It’s not limited to the algorithmic or statistical aspects.

Here are some fields data science covers:

* data integration
* distributed architecture
* data visualization
* data engineering
* deployment in production mode
* data-driven decisions

So while ML experts are busy with building useful algorithms throughout the project lifecycle, data scientists have to be more flexible switching between different data roles according to the needs of the project.

**2.WHY IS DATA SCIENCE IS REQUIRED FOR MACHINE LEARNING ENTHUSIASTS?**

Machine Learning done means you are half way though data science. So, some part of ML is mandatory to learn data science. The overlap between these two fields is enormous. You cannot use machine learning algorithms with real world data, with-out data processing/data-science—-as real world data almost never comes in a nice labeled structured format that algorithms can understand. And if your doing “data science”…even most data visualization algorithms are in truth what is considered ***machine learning algorithms***.

Outside of job titles/descriptions, in my book data science and machine learning are basically two sides to the same field.

**3.MODELS PRESENT IN MACHINE LEARNING ALGORITHMS**

**1. Supervised Learning**

**How it works:** This algorithm consist of a target / outcome variable (or dependent variable) which is to be predicted from a given set of predictors (independent variables). Using these set of variables, we generate a function that map inputs to desired outputs. The training process continues until the model achieves a desired level of accuracy on the training data. Examples of Supervised Learning: Regression, [Decision Tree](https://www.analyticsvidhya.com/blog/2015/01/decision-tree-simplified/), [Random Forest](https://www.analyticsvidhya.com/blog/2014/06/introduction-random-forest-simplified/), KNN, Logistic Regression etc.

**2. Unsupervised Learning**

**How it works:**In this algorithm, we do not have any target or outcome variable to predict / estimate. It is used for clustering population in different groups, which is widely used for segmenting customers in different groups for specific intervention. Examples of Unsupervised Learning: Apriori algorithm, K-means.

**3. Reinforcement Learning:**

**How it works:** Using this algorithm, the machine is trained to make specific decisions. It works this way: the machine is exposed to an environment where it trains itself continually using trial and error. This machine learns from past experience and tries to capture the best possible knowledge to make accurate business decisions. Example of Reinforcement Learning: Markov Decision Process

**Others:**

**1. Regression Algorithms**

* Ordinary Least Squares Regression (OLSR)
* Linear Regression
* Logistic Regression
* Stepwise Regression
* Multivariate Adaptive Regression Splines (MARS)
* Locally Estimated Scatterplot Smoothing (LOESS)

**2. Instance-based Algorithms**

* k-Nearest Neighbour (kNN)
* Learning Vector Quantization (LVQ)
* Self-Organizing Map (SOM)
* Locally Weighted Learning (LWL)

**3. Regularization Algorithms**

* Ridge Regression
* Least Absolute Shrinkage and Selection Operator (LASSO)
* Elastic Net
* Least-Angle Regression (LARS)

**4. Decision Tree Algorithms**

* Classification and Regression Tree (CART)
* Iterative Dichotomiser 3 (ID3)
* C4.5 and C5.0 (different versions of a powerful approach)
* Chi-squared Automatic Interaction Detection (CHAID)
* Decision Stump
* M5
* Conditional Decision Trees

**5. Bayesian Algorithms**

* Naive Bayes
* Gaussian Naive Bayes
* Multinomial Naive Bayes
* Averaged One-Dependence Estimators (AODE)
* Bayesian Belief Network (BBN)
* Bayesian Network (BN)

**6. Clustering Algorithms**

* k-Means
* k-Medians
* Expectation Maximisation (EM)
* Hierarchical Clustering

**7. Association Rule Learning Algorithms**

* Apriori algorithm
* Eclat algorithm

**8. Artificial Neural Network Algorithms**

* Perceptron
* Back-Propagation
* Hopfield Network
* Radial Basis Function Network (RBFN)

**9. Deep Learning Algorithms**

* Deep Boltzmann Machine (DBM)
* Deep Belief Networks (DBN)
* Convolutional Neural Network (CNN)
* Stacked Auto-Encoders

**10. Dimensionality Reduction Algorithms**

* Principal Component Analysis (PCA)
* Principal Component Regression (PCR)
* Partial Least Squares Regression (PLSR)
* Sammon Mapping
* Multidimensional Scaling (MDS)
* Projection Pursuit
* Linear Discriminant Analysis (LDA)
* Mixture Discriminant Analysis (MDA)
* Quadratic Discriminant Analysis (QDA)
* Flexible Discriminant Analysis (FDA)

**11. Ensemble Algorithms**

* Boosting
* Bootstrapped Aggregation (Bagging)
* AdaBoost
* Stacked Generalization (blending)
* Gradient Boosting Machines (GBM)
* Gradient Boosted Regression Trees (GBRT)
* Random Forest

**12. Other Algorithms**

* Computational intelligence (evolutionary algorithms, etc.)
* Computer Vision (CV)
* Natural Language Processing (NLP)
* Recommender Systems
* Reinforcement Learning
* Graphical Models