A Brief Review of Key Machine Learning Algorithms, Models Explainability and Business Use Cases

Source and Credits: https://www.mckinsey.com/

https://www.linkedin.com/in/ratnakarpandey/

Supervised ML Models

https://www.linkedin.com/in/ratnakarpandey/

Linear regression

Highly interpretable, standard method for modeling the past relationship between independent input variables and dependent output variables (which can have an infinite number of values) to help predict future values of the output variables

Business use cases

Understand product-sales drivers such as competition prices, distribution, advertisement, etc

Optimize price points and estimate product-price elasticities

Logistic regression

A model with some similarities to linear regression that's used for classification tasks, meaning the output variable is binary (eg, only black or white) rather than continuous (eg, an infinite list of potential colors)

Business use cases

Classify customers based on how likely they are to repay a loan

Predict if a skin lesion is benign or malignant based on its characteristics (size, shape, color, etc)

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Linear/quadratic discriminant analysis

Upgrades a logistic regression to deal with nonlinear problems–those in which changes to the value of input variables do not result in proportional changes to the output variables

Business use cases

Predict client churn

Predict a sales lead's likelihood of closing

Decision tree

Highly interpretable classification or regression model that splits data-feature values into branches at decision nodes (eg, if a feature is a color, each possible color becomes a new branch) until a final decision output is made

Business use cases

Understand product attributes that make a product most likely to be purchased

Provide a decision framework for hiring new employees

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Naive Bayes

Classification technique that applies Bayes theorem, which allows the probability of an event to be calculated based on knowledge of factors that might affect that event (eg, if an email contains the word "money," then the probability of it being spam is high)

Business use cases

Analyze sentiment to assess product perception in the market

Create classifiers to filter spam emails

Support vector machine

A technique that's typically used for classification but can be transformed to perform regression. It draws a division between classes that's as wide as possible. It also can be generalized to solve nonlinear problems.

Business use cases

Predict how many patients a hospital will need to serve in a time period

Predict how likely someone is to click on an online ad

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Random forest

Classification or regression model that improves the accuracy of a simple decision tree by generating multiple decision trees and taking a majority vote of them to predict the output, which is a continuous variable (eg, age) for a regression problem and a discrete variable (eg, either black, white, or red) for classifcation

Business use cases

Predict call volume in call centers for staffing decisions

Predict power usage in an electrical-distribution grid

AdaBoost

Classification or regression technique that uses a multitude of models to come up with a decision but weighs them based on their accuracy in predicting the outcome

Business use cases

Detect fraudulent activity in credit-card transactions. Achieves lower accuracy than deep learning

Simple, low-cost way to classify images (eg, recognize land usage from satellite images for climate-change models). Achieves lower accuracy than deep learning

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Gradient-boosting trees

Classification or regression technique that generates decision trees sequentially, where each tree focuses on correcting the errors coming from the previous tree model. The final output is a combination of the results from all trees

Business use cases

Forecast product demand and inventory levels

Predict the price of cars based on their characteristics (eg, age and mileage)

Simple neural network

Model in which artificial neurons (software-based calculators) make up an input layer, one or more hidden layers where calculations take place, and an output layer. It can be used to classify data or find the relationship between variables in regression problems.

Business use cases

Predict the probability that a patient joins a healthcare program

Predict whether registered users will be willing or not to pay a particular price for a product

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Un-Supervised ML Models

https://www.linkedin.com/in/ratnakarpandey/

Autoencoder

A type of neural network that can be used to represent data efficiently by removing unnecessary information. The representation can be thought of as a compressed version of the original data.

Business use cases

Reduce noise in a medical image (eg, MRI) to analyze it more accurately

Detect fake reviews and opinions on social media

K-means clustering

Puts data into a number of groups (k) that each contain data with similar characteristics (as determined by the model, not in advance by humans)

Business use cases

Segment customers into groups by distinct characteristics (eg, age group)–for instance, to better assign marketing campaigns or prevent churn

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Gaussian mixture model

A generalization of k-means clustering that provides more flexibility in the size and shape of groups (clusters)

Business use cases

Segment employees based on likelihood of attrition

Segment customers to better assign marketing campaigns using less-distinct customer characteristics (eg, product preferences)

Hierarchical clustering

Splits or aggregates clusters along a hierarchical tree to form a classification system

Business use cases

Cluster loyalty-card customers into progressively more microsegmented groups

Inform product usage/development by grouping customers mentioning keywords in social-media data

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Recommender system

Often uses cluster behavior prediction to identify the important data necessary for making a recommendation

Business use cases

Recommend what movies consumers should view based on preferences of other customers with similar attributes

Recommend news articles a reader might want to read based on the article she or he is reading

Manifold learning

An approach to representing data with high dimensionality in a two-dimensional chart to more easily visualize interesting patterns in the data

Business use cases

Detect patterns in spread of a pandemic

Group similar customers together and recommend next best product for them to buy

Predict brain-tumor progression

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Models Explainability

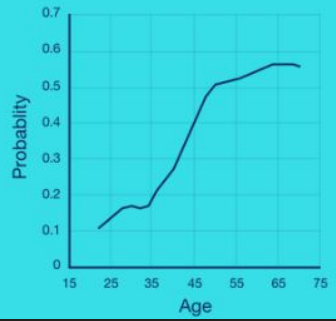
https://www.linkedin.com/in/ratnakarpandey/

Partial dependence plots (PDP)

Plots that trace how the output changes as a variable is changed.

Example

Generally, as age increases, the probability of income being high increases, with a steep increase between ages 35 to 50 (with all other features at their average values).

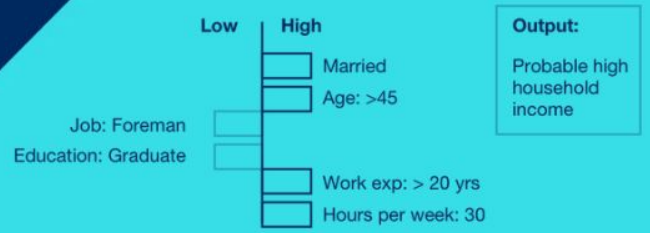


Local interpretable model-agnostic explanations (LIME)

LIME studies how variations of input data affect the output, enabling it to list features affecting the output and score their level of impact.

Example

The probability of a high income is driven by marital status, age, years of work experience, and hours worked per week, while the probability of income being low is driven by occupation and education.



https://www.linkedin.com/in/ratnakarpandey/

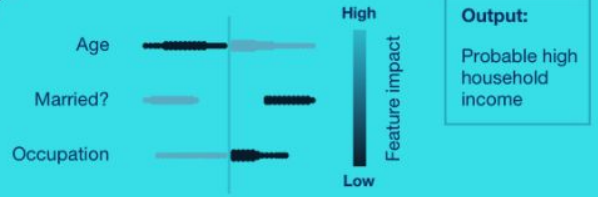
Source and Credits: https://www.mckinsey.com/

SHapley Additive exPlanations (SHAP)

SHAP uses a game theory approach to score each feature based on its contribution to the output after considering the feature's interaction with all other features. Unlike LIME, SHAP offers a global explanation of features by considering all possible feature interactions as well as individual observation-level explanations.

Example

The Y axis shows the relative importance of each feature to income level, while the X axis shows the correlation to income level (eg, age has a high impact on predicting high income and high age correlates with high income).



Output:

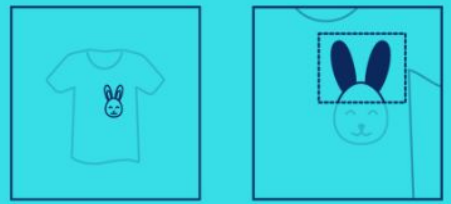
Probable high household income

Saliency Maps

Heat maps, primarily used for image-classification models, in which the degree of heat in a region on the image corresponds to the level of impact that region has on predicting what the image shows.

Example

The regions highlighted in dark blue on the right-hand image are the pixels contributing most to this image being classified as a rabbit.



Image

Saliency map

https://www.linkedin.com/in/ratnakarpandey/

Source and Credits: https://www.mckinsey.com/

Hope you like it! Do share your reactions, comments.

https://www.linkedin.com/in/ratnakarpandey/