OneModel Academy

One AI

Regressions

Machine Learning

Hayley Bresina
One AI Client Enablement

07 Jun 2024

Topics Covered

- An overview of regressions
- Common regression algorithms
- Strengths & weaknesses of regressions
- People analytics use cases

Learning Outcomes

You will:

- Understand regressions as a supervised machine learning technique for predicting continuous outputs
- Be introduced to common regression algorithms such as linear regression, ridge, lasso, elastic net, & random forest
- Identify how One AI takes advantage of the strengths & mitigates the weaknesses of regressions
- Explore regression model use cases & how these models enable organizations to extract valuable insights from employee data & inform decision-making

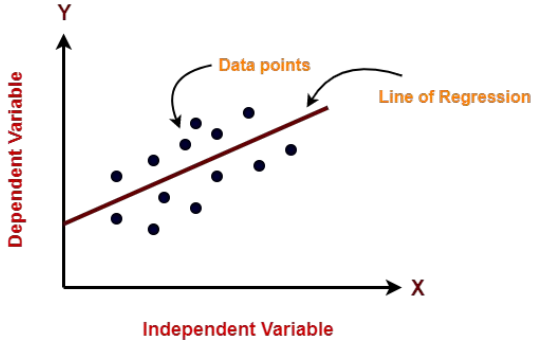
Regression Overview

Overview

- **Regression** is a supervised ML technique used for predicting continuous values based on learning relationships from historical data
  - Goal: **accurately** predict a target value based on input features

Target variable = dependant variable

Input feature = independent variable



  - Learns by analyzing the **relationships** between the input features & target values that it observes from each instance in the **historical data**
  - **Output:** predicts a continuous value by plotting a line or curve of best fit

Overview (Cont)

Types

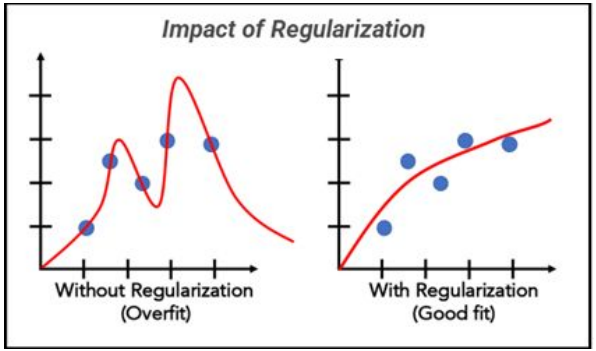
- **Linear**: 1 input & 1 output
- **Multiple**: many inputs & 1 output
- **Multivariate**: many inputs & many outputs

Training

- **Labeled dataset is required** - datasets containing examples where both the input features and the corresponding target values are known
  - Performance rating example:
    - Performance rating = target value
    - Data points like job role, education, skills, location, etc. = input features

Regularization

- **Regularization** is a technique used to **prevent overfitting** & improve the generalization ability of regression models
  - Methods introduce a penalty term that encourage the model to prefer **simpler patterns** that are more likely to generalize well to new, unseen data
  - Goal: strike a **balance** between fitting the training data well & avoiding excessive complexity



Regression Algorithms

Regression Algorithms

- Machine learning use **algorithms** to learn from data, identify patterns, make predictions, or perform tasks without explicit programming
  - An **algorithm** is the mathematical procedure, technique, or set of rules that the model follows to do so
- Common regression algorithms
  - Linear regression
  - Ridge, Lasso, & Elastic Net
  - Decision tree
  - Random forest
  - Gradient boosting machines (GBM)

Strengths & Weaknesses

Strengths & Weaknesses

Strengths

- Interpretability
- Simplicity & speed
- Versatility
- Feature selection & feature importance

Weaknesses

- Overfitting
- Limited capacity to capture complex non-linear relationships
- Non-robustness to missing data
- Data requirements (labeled data)

People Analytics Use Cases

People Analytics Use Cases

- Extract valuable insights from employee & recruiting data
- Go beyond the predictions & uncover critical relationships & factors influencing workforce dynamics
    - Understand the impact of various factors on employee outcomes, such as productivity, compensation satisfaction, or job engagement
- Possible regression models:
    - Group or employee attrition
    - Salary & compensation
    - Engagement
    - Custom models

Thanks for watching!