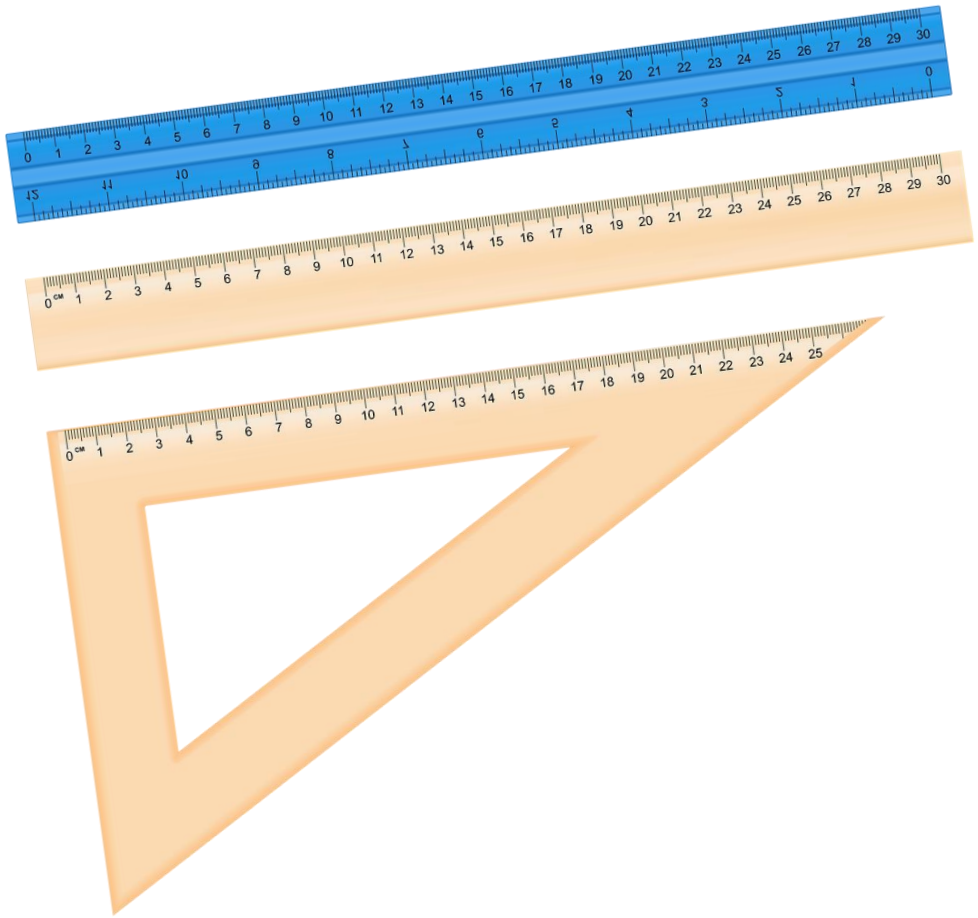
Train In Data

Feature Scaling

Feature Magnitude matters

- The regression coefficient is directly influenced by the scale of the variable
- Variables with bigger magnitude / value range dominate over the ones with smaller magnitude / value range
- Gradient descent converges faster when features are on similar scales
- Feature scaling helps decrease the time to find support vectors for SVMs
- Euclidean distances are sensitive to feature magnitude.

Algorithms sensitive to magnitude



The machine learning models affected by the magnitude of the feature:

- Linear and Logistic Regression
- Neural Networks
- Support Vector Machines
- KNN
- K-means clustering
- Linear Discriminant Analysis (LDA)
- Principal Component Analysis (PCA)

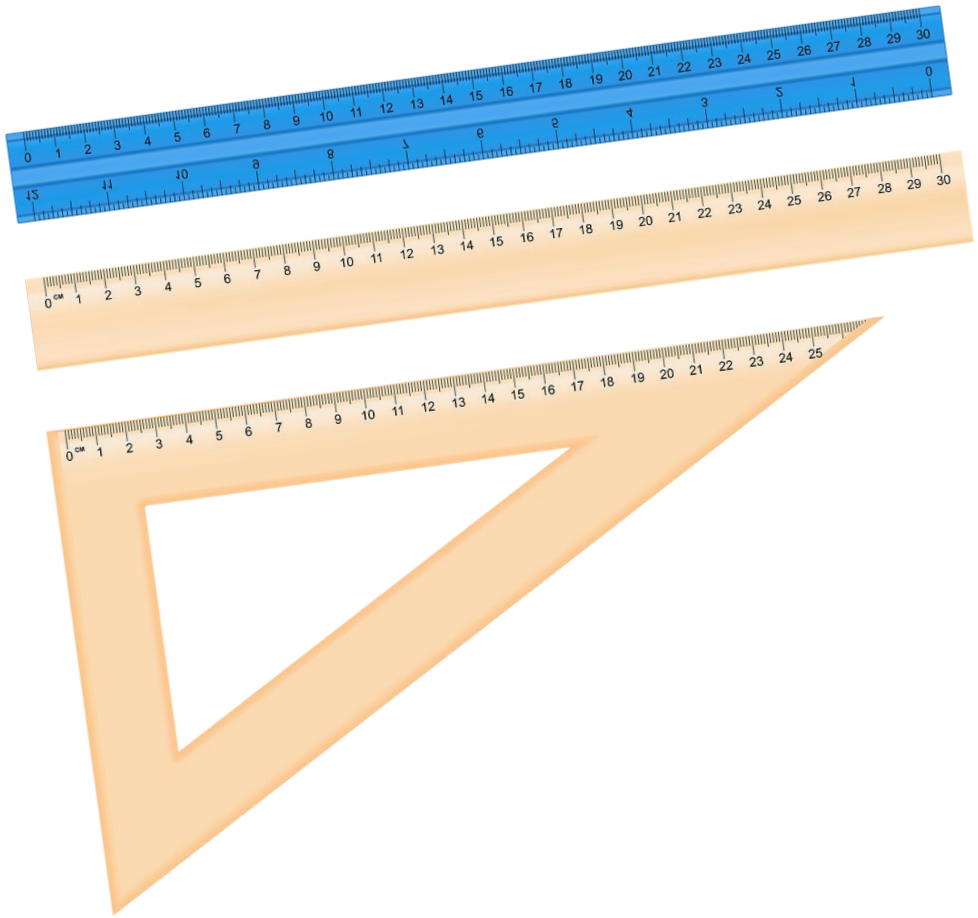
Machine learning models insensitive to feature magnitude are the ones based on Trees:

- Classification and Regression Trees
- Random Forests
- Gradient Boosted Trees

Feature Scaling

- Feature scaling refers to the methods used to normalize the range of values of independent variables.
- In other words, the methods to set the feature value range within a similar scale.
- Feature scaling is generally the last step in the data pre-processing pipeline, performed just before training the machine learning algorithms.

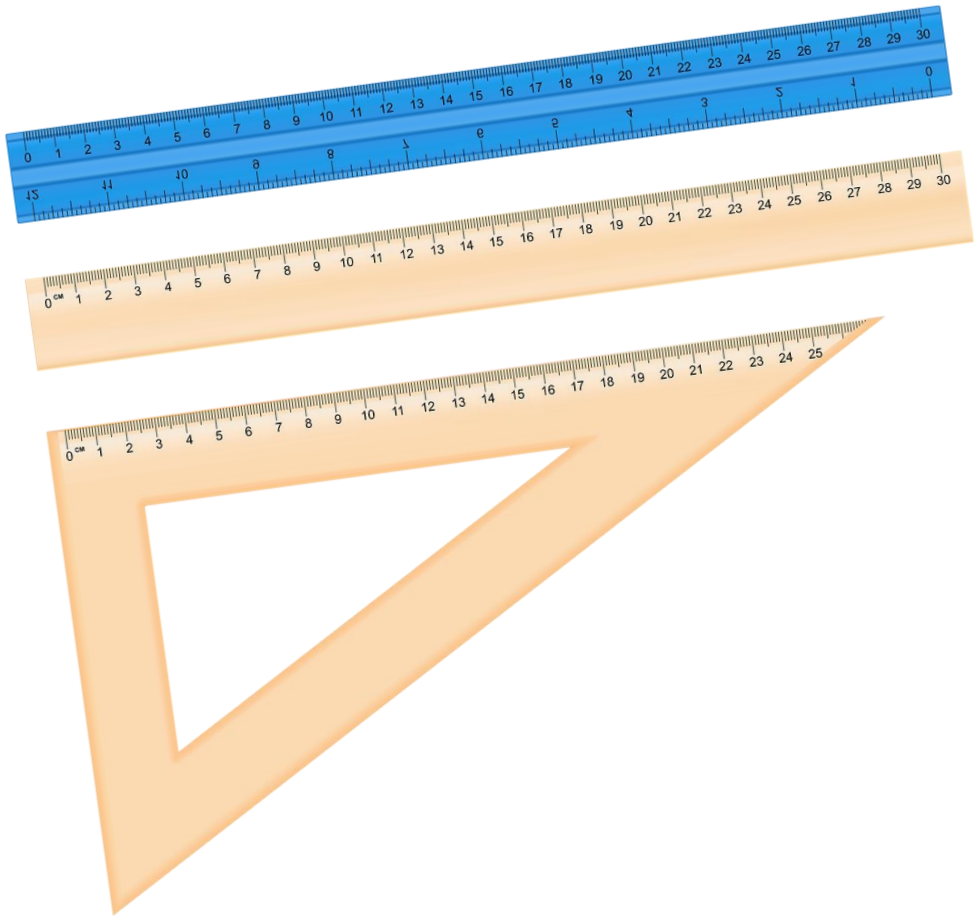
Feature scaling methods



Scaling methods

- Standardisation
- Mean normalisation
- Scaling to maximum and minimum
- Scaling to absolute maximum
- Scaling to median and quantiles
- Scaling to unit norm

Feature scaling methods



Scaling methods

- **Standardisation**
- Mean normalisation
- **Scaling to maximum and minimum**
- Scaling to absolute maximum
- Scaling to median and quantiles
- Scaling to unit norm

Accompanying Jupyter Notebook

- Read the accompanying Jupyter Notebook
- Demo on how to implement feature selection algorithms using Scikit-learn

Train In Data

THANK YOU

www.trainindata.com