



Dear participant,

Welcome to the study on understanding the explainability and comprehensibility of different machine learning models.

Parts of our study are implemented with JavaScript. If deactivated, we kindly recommend to reactivate it for the study. Furthermore, please use full-screen mode.

Approx. time: 15 min.

You can start the survey by clicking on ">>" at the bottom-right of this page.



Lukas-Valentin
Herm



Jonas
Wanner



Prof. Dr.
Christian Janiesch



Franz Seubert



Prof. Dr.
Axel Winkelmann

**Please specify your gender.**

- ☐ male
- ☐ female
- ☐ diverse
- ☐ no answer

Please specify your age.

- ☐ <20
- ☐ 20 - 30
- ☐ 31 - 40
- ☐ 41 - 50
- ☐ 51 - 60
- ☐ >60
- ☐ no answer

Please specify your location.

- ☐ Africa
- ☐ Asia
- ☐ Europe
- ☐ North America
- ☐ South America
- ☐ Oceania

Please specify your experience with artificial intelligence.

- ☐ none
- ☐ < 2 yrs
- ☐ 2 - 5 yrs
- ☐ 6 - 10 yrs
- ☐ > 10 yrs

Please specify your general degree of trust in artificial intelligence solutions.

very low low medium high very high



Please specify your willingness to accept artificial intelligence at your workplace.

very low low medium high very high



Please specify your willingness to take risks.

very low low medium high very high



Lukas-Valentin
Herm



Jonas
Wanner



Prof. Dr.
Christian Janiesch



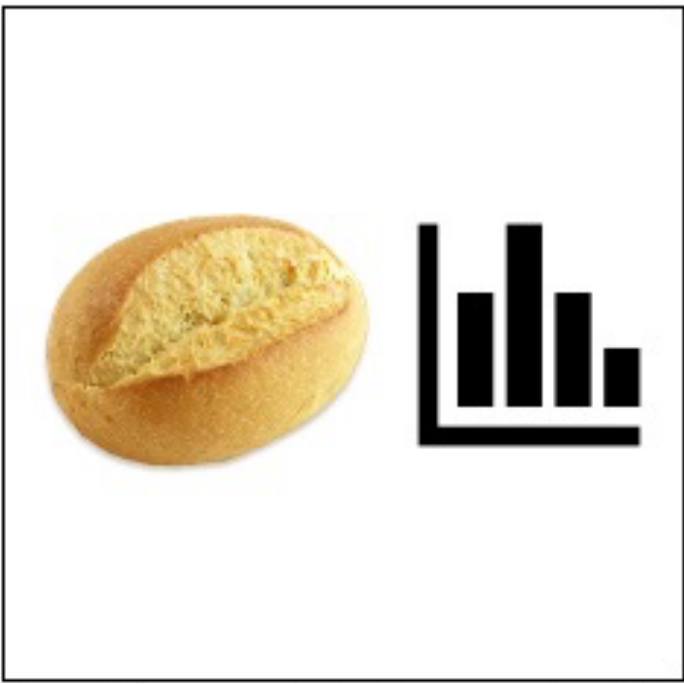
Franz Seubert



Prof. Dr.
Axel Winkelmann

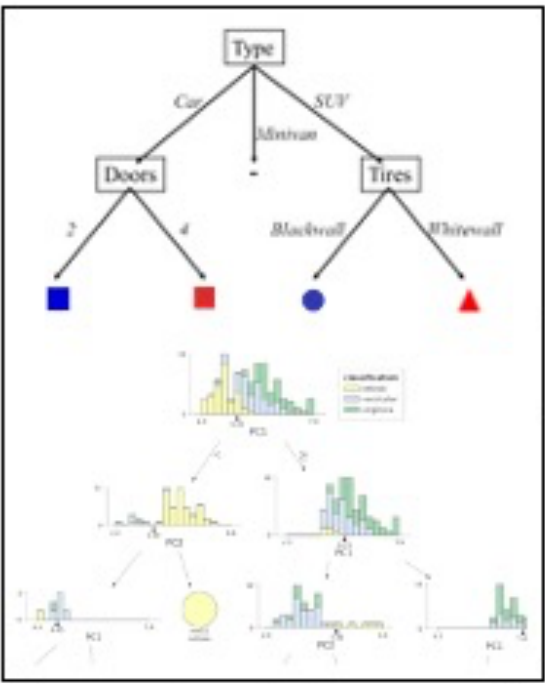
What are we going to do?

1. Presentation of Use Cases



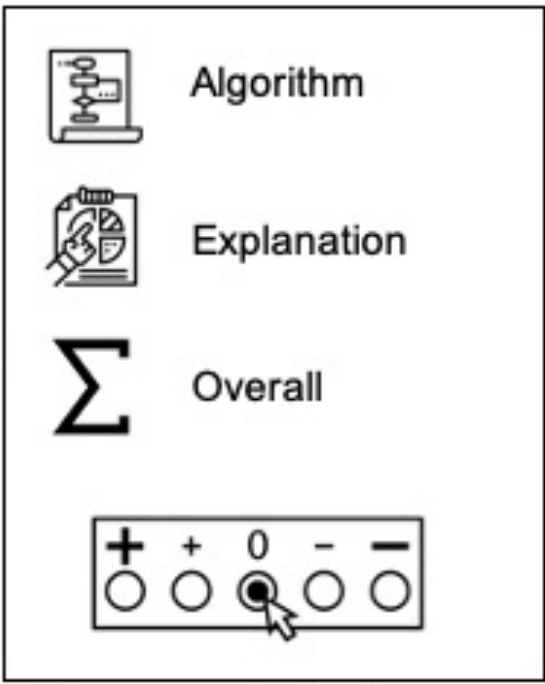
You will be presented with **one** real world **use case**, where you need to decide for an AI-based decision support system to assist you.

2. Presentation of ML Types



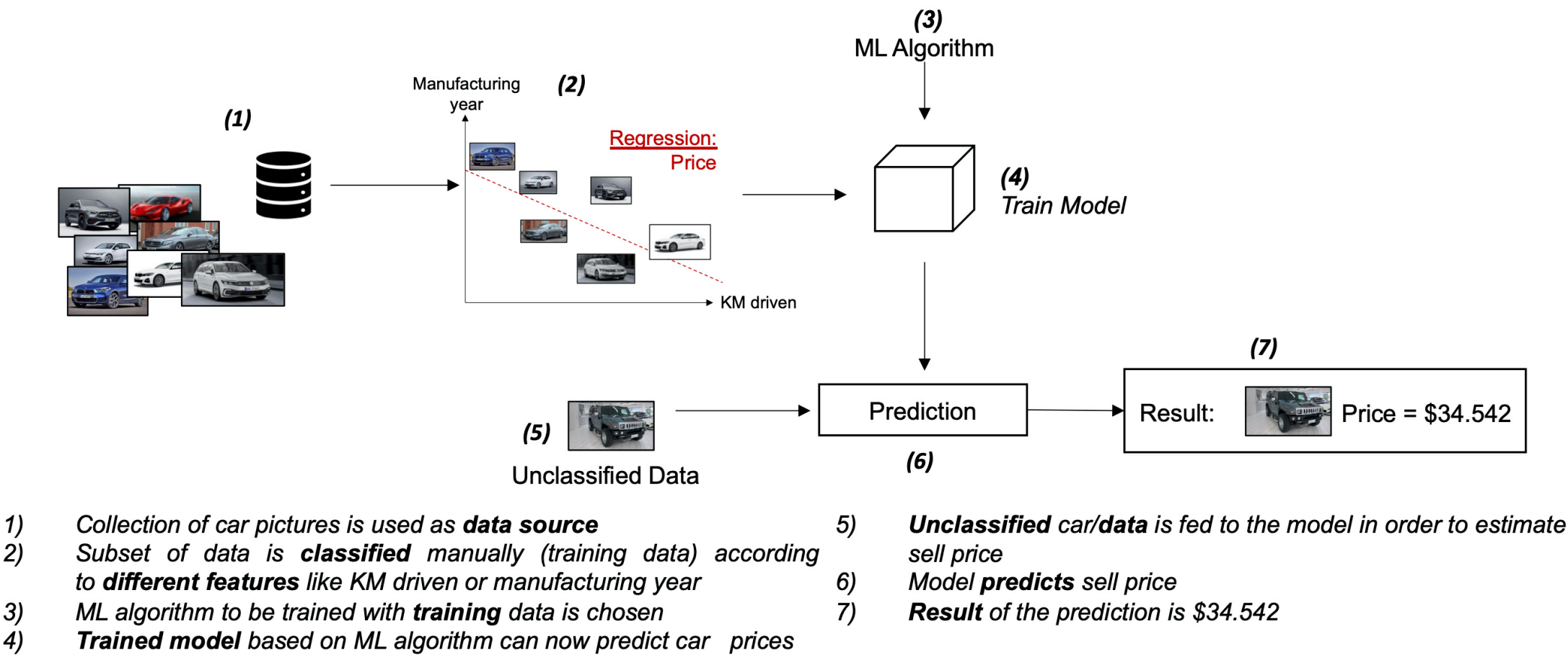
To support your decision, you will receive an introduction to each AI-based system with its **algorithm** and **result presentation**.

3. Perception of ML Types



Each use case is about **your perception** of each ML model's explanatory goodness.

What is „Machine Learning“ (ML)?



What to know about ML-based explanations?

Important information about ML-based explanations

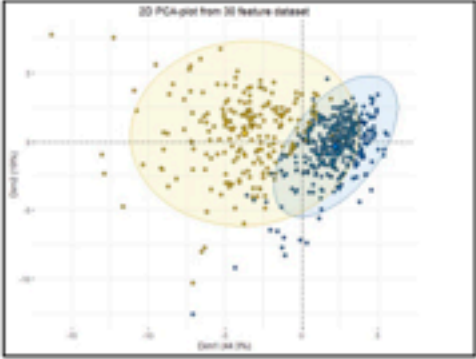
1) Complex dataset

id	cycle	setting1	setting2	setting3	s1	s2	s3	s4	s5	...	s13	s14	s15	s16	s17	s18	s19	s20	s21	RUL
1	1	0.458770	0.166667	0.0	0.0	0.183735	0.406802	0.309757	0.0	...	0.205882	0.199608	0.363986	0.0	0.333333	0.0	0.0	0.713178	0.724662	191
1	2	0.609195	0.250000	0.0	0.0	0.283133	0.453019	0.352633	0.0	...	0.279412	0.162813	0.411312	0.0	0.333333	0.0	0.0	0.666667	0.731014	190
1	3	0.252874	0.750000	0.0	0.0	0.343373	0.369523	0.370527	0.0	...	0.220588	0.171793	0.357445	0.0	0.166667	0.0	0.0	0.827907	0.621375	189
1	4	0.540230	0.500000	0.0	0.0	0.343373	0.256159	0.331195	0.0	...	0.294118	0.174889	0.166603	0.0	0.333333	0.0	0.0	0.573643	0.662386	188
1	5	0.390805	0.333333	0.0	0.0	0.348398	0.257467	0.404625	0.0	...	0.235294	0.174734	0.402078	0.0	0.416667	0.0	0.0	0.589147	0.704502	187
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
100	196	0.477011	0.250000	0.0	0.0	0.686747	0.587312	0.782917	0.0	...	0.558824	0.194344	0.656791	0.0	0.750000	0.0	0.0	0.271318	0.109500	4
100	197	0.408046	0.083333	0.0	0.0	0.701807	0.729453	0.866475	0.0	...	0.500000	0.188668	0.727203	0.0	0.583333	0.0	0.0	0.124031	0.366197	3

2) Pre-processing

3) Training & Prediction

4) Visualization



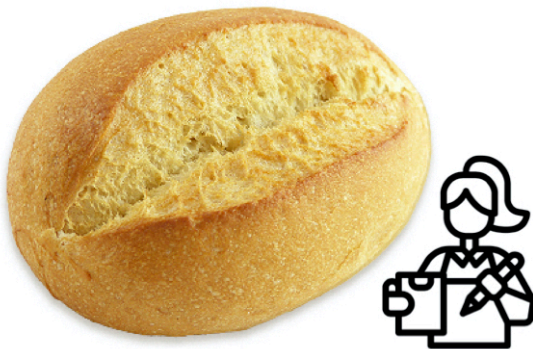
1) ML models analyze **complex issues** in **high-dimensional datasets**

2) Normally a **cleansing** of the **data** set takes place. For example, this involves replacing unrealistic or missing values.

3) The pre-processed dataset is used for **training ML models** as well as for predicting new results

4) The **resulting visualization** is a plot of the model's decision logic for the respective case

Use Case



Imagine you are a baker and have to deal with the following problem:
Every day you have to decide **how many buns you want to order** for the sale on the next day. In order to make the task easier, **AI-based models** should be introduced in your company. To help your supervisor making the right choice, he or she will ask for your feedback on the various AI models.

Each model was provided with the *delivery, returns and sales features as well as weather and calendar data of the last three years*.

Weekday (numeric)	Month (numeric)	School holidays (binary)	Holidays (binary)	Next 2 days holidays (binary)	Rainfall in mm	Temperature in Celsius	Average demand weekday (last 4 weeks)	Average growth (comparison last week)	Relative Store size	Advertising in current week (binary)	Advertising in last week binary	Actual Demand
1	1	1	1	0	11.02	31.20	34	1.23	0.52	1	0	156.2
2	2	0	1	0	10.56	34.00	234	1.51	0.56	0	0	185.9
3	7	1	0	1	11.00	34.89	345	1.53	0.62	1	0	236.3

Through this data, the various models now generate **order proposals** for the next day. The **explanations** of the models show how the *different features* influence the **sales quantity** of the buns.

Goal: Determine the optimal order quantity.
Setting: medium-stake, costs for wrong decision

Presentation of Study Design

In the following you will be given information on **different ML** models. You should rate them regarding their **explanatory goodness**. Assume a uniform performance of all ML models.

Machine Learning Model

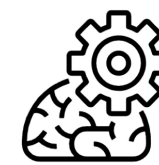
Algorithm Explanation

Decision Explanation


 *graphical result*

The presented explanations are good.

Strongly disagree
Disagree
More or less disagree
Undecided
More or less agree
Agree
Strongly Agree



Algorithm Explanation: The explanation of the heart of the ML model – it is the logical foundation for its decision making.



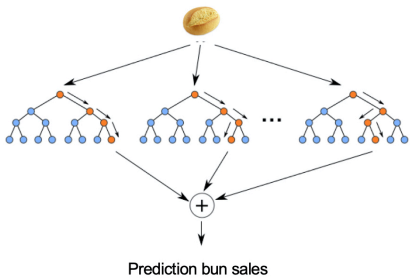
Decision Explanation: The ML model's (graphical) explanation for the respective case. In many cases, a mouse-over on the image can be used to zoom in.



Your perception to the explanatory goodness of the model

Please provide your answers only with respect to the presented use case!

General Introduction: Algorithm of Random Forest

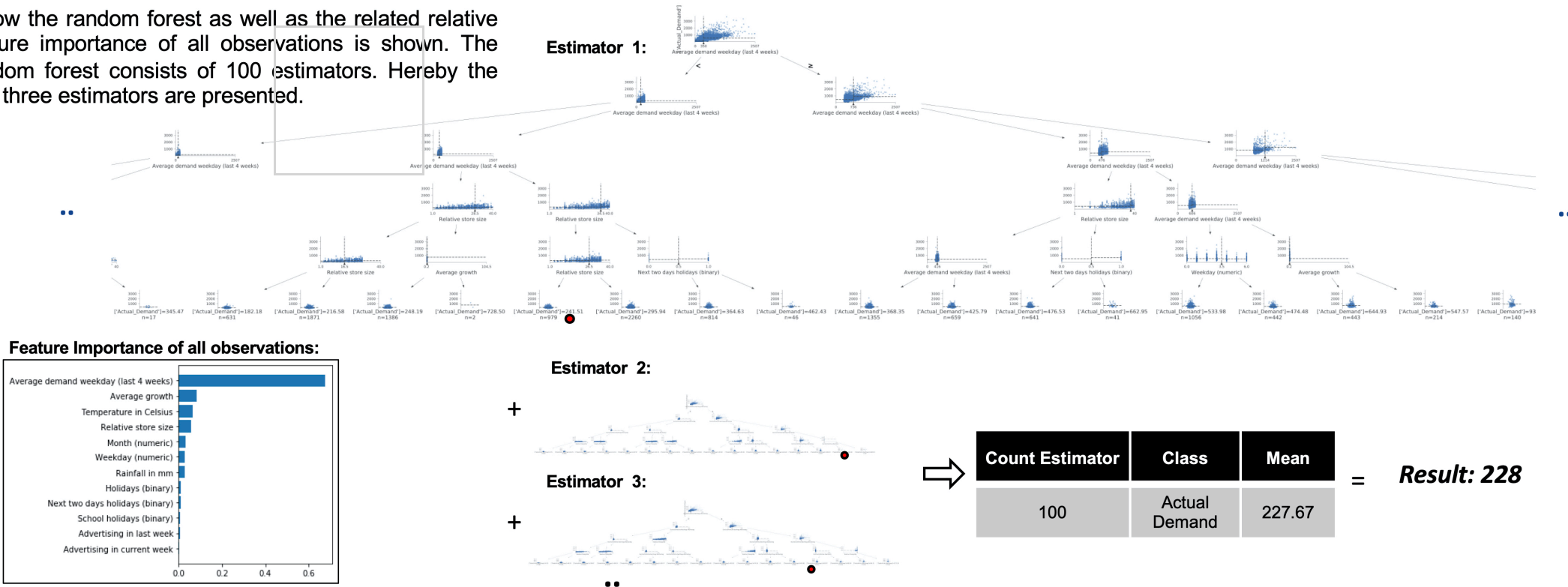


A random forest is a method consisting of **several uncorrelated decision trees also called estimators**. A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**. The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases. All decision trees have grown under a certain type of randomization during the learning process.

For a prediction, each tree in that forest makes a decision, according to the features, and the **mean of all decision** determines the final result. In the **example** on the left side, the output would be the prediction of bun sales for tomorrow.

Result: Bun Order Quantity

Below the random forest as well as the related relative feature importance of all observations is shown. The random forest consists of 100 estimators. Hereby the first three estimators are presented.



Rate the following statement: "The presented explanations are good."

Please consider both, your perceived explanatory goodness of the algorithm's decision logic and its result.

Strongly disagree

Disagree

More or less disagree

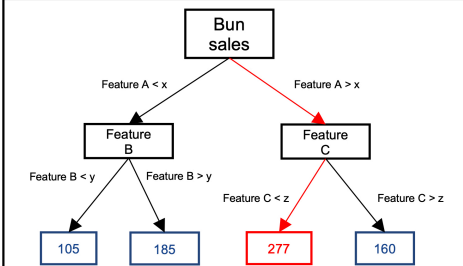
Undecided

More or less agree

Agree

Strongly Agree

General Introduction: Algorithm of Decision Trees



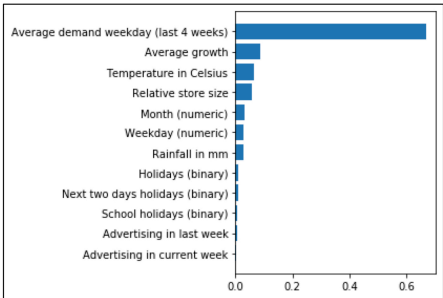
A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**.

The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases.

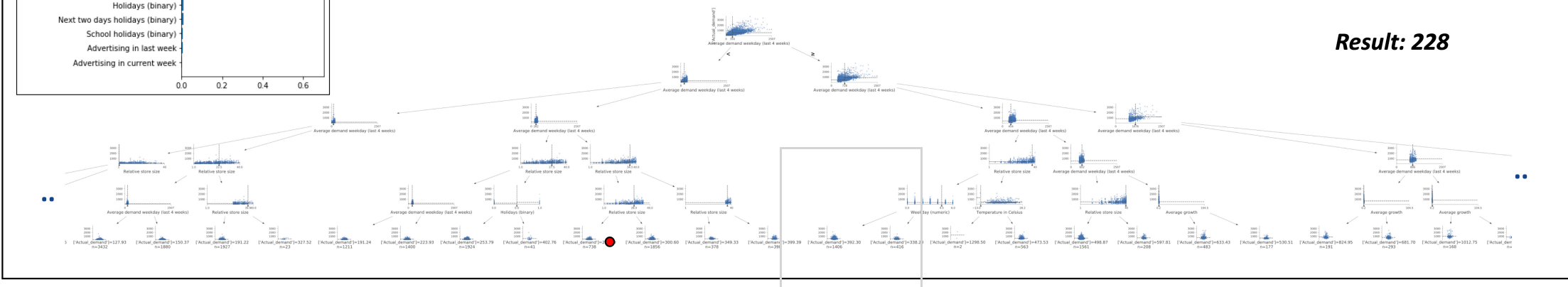
In the **example** on the left side, the output of the sales data would be 277 according to feature A and C (marked red).

Result: Bun Order Quantity

Feature Importance of all observations:



The decision tree as well as the related relative feature importance of all observations is shown. The features of the inserted data are used to predict the result.



Rate the following statement: "The presented explanations are good."

Please consider both, your perceived explanatory goodness of the algorithm's decision logic and its result.

Strongly disagree

Disagree

More or less disagree

Undecided

More or less agree

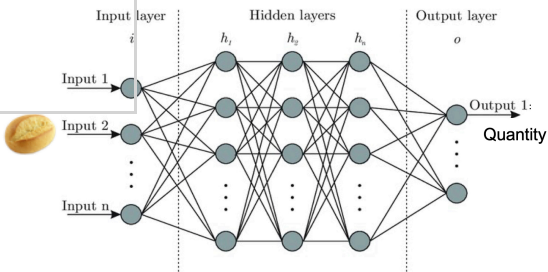
Agree

Strongly Agree



44%

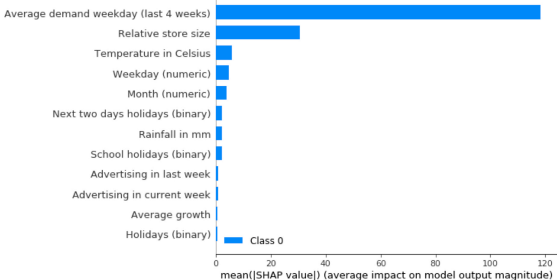
General Introduction: Algorithm of Artificial Neural Networks (ANN)



An artificial neural network (ANN) is an information processing paradigm that is inspired by the way the **biological nervous system** such as a human brain processes information. The algorithm is composed of a large number of highly **interconnected processing elements** (neurons) working in unison to solve a specific problem. The input layer receives new cases, which are calculated, through different mathematical operations, within the hidden layers. In the **example** on the left side sales data is forwarded to the ANN. Using many calculations in different layers, the model estimate the upcoming sales for buns.

Result: Bun Order Quantity

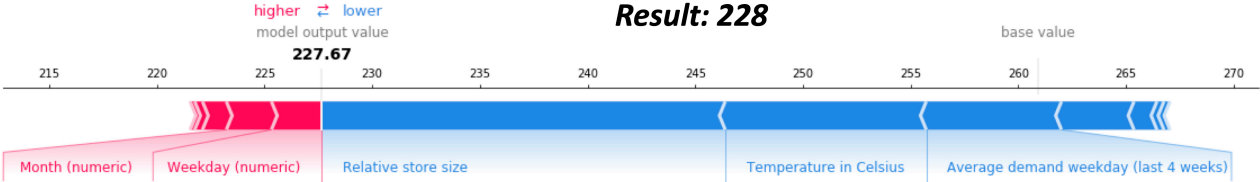
Feature importance of all observations:



On the left side, the mean global influence of the different features within the dataset in percent is shown. On the right side, the features with the influence (positive: red, negative: blue) on a specific prediction (model output value) based on the mean of all predictions (base value) in the dataset are presented.

Result: 228

base value



Rate the following statement: "The presented explanations are good."

Please consider both, your perceived explanatory goodness of the algorithm's decision logic and its result.

Strongly disagree

Disagree

More or less disagree

Undecided

More or less agree

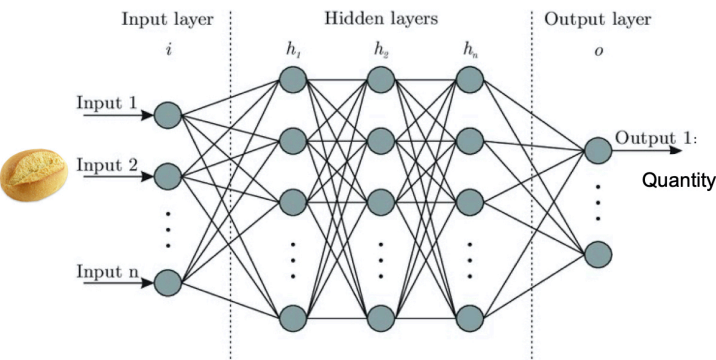
Agree

Strongly Agree

<<

>>

General Introduction: Algorithm of Artificial Neural Networks (ANN)



An artificial neural network (ANN) is an information processing paradigm that is inspired by the way the **biological nervous system** such as a human brain processes information.

The algorithm is composed of a large number of highly **interconnected processing elements** (neurons) working in unison to solve a specific problem. The input layer receives new cases, which are calculated, through different mathematical operations, within the hidden layers.

In the **example** on the left side sales data is forwarded to the ANN. Using many calculations in different layers, the model estimate the upcoming sales for buns. The processing of the hidden layers is typically neither accessible nor helpful to the user.

Result: Bun Order Quantity

*Black box:
no visualization available*

The model predicts 228 without any explanation.

Result: 228

Rate the following statement: "The presented explanations are good."

Please consider both, your perceived explanatory goodness of the algorithm's decision logic and its result.

Strongly disagree

Disagree

More or less disagree

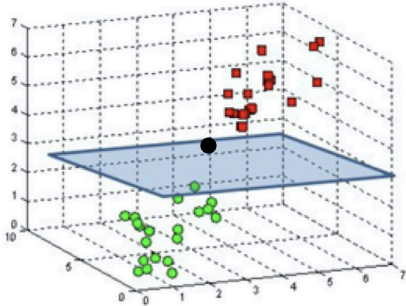
Undecided

More or less agree

Agree

Strongly Agree

General Introduction: Algorithm of Support Vector Machines (SVM)



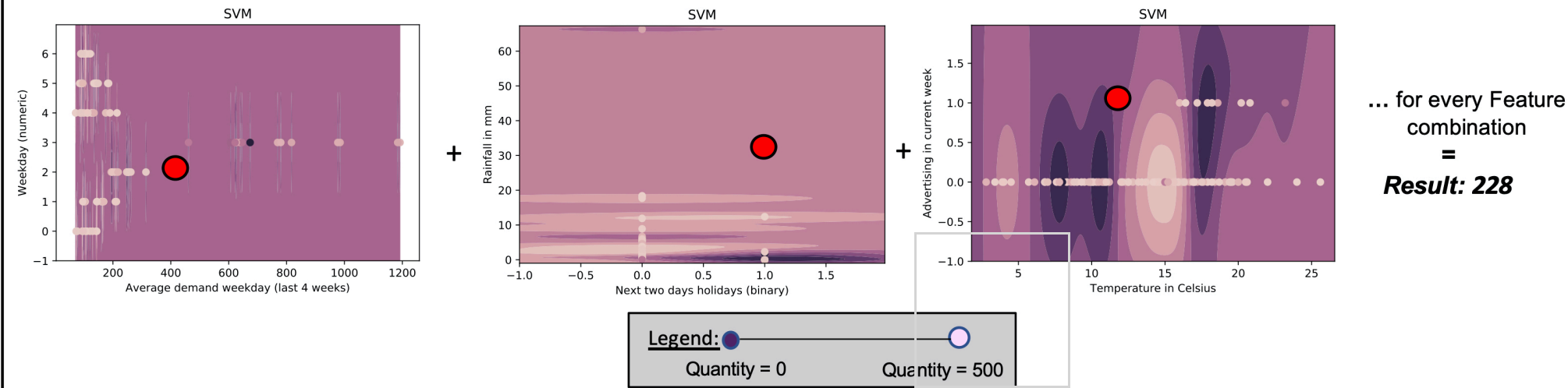
The SVM algorithm receives **data with n-features**, which are **represented as a n-dimensional vector**, whereby the individual features are assigned to a class.

Based on the vector space, the algorithm determines the **optimum dividing hyperplanes (blue) between the features**, where the distance (margin) to the nearest points of the other class should be maximized. This wide, empty border later ensure that even objects that do not correspond exactly to the training objects are classified as reliably as possible.

The hyperplanes classify the sales of buns according to its various features. When a particular observation is inserted into the SVM, the SVM predicts the sales of buns for tomorrow.

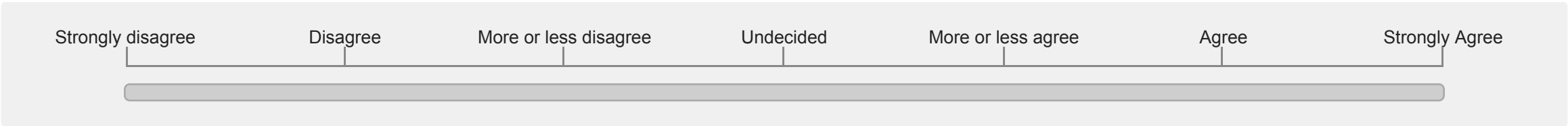
Result: Bun Order Quantity

In order so show the explanation for the decision, multiple pairwise comparisons of two features according to the bun order quantity is shown below. This is done for every possible feature combination.

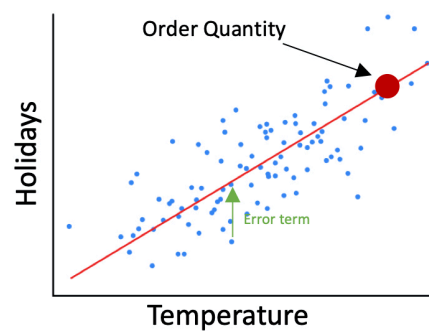


Rate the following statement: "The presented explanations are good."

Please consider both, your perceived explanatory goodness of the algorithm's decision logic and its result.



General Introduction: Algorithm of Linear Regression



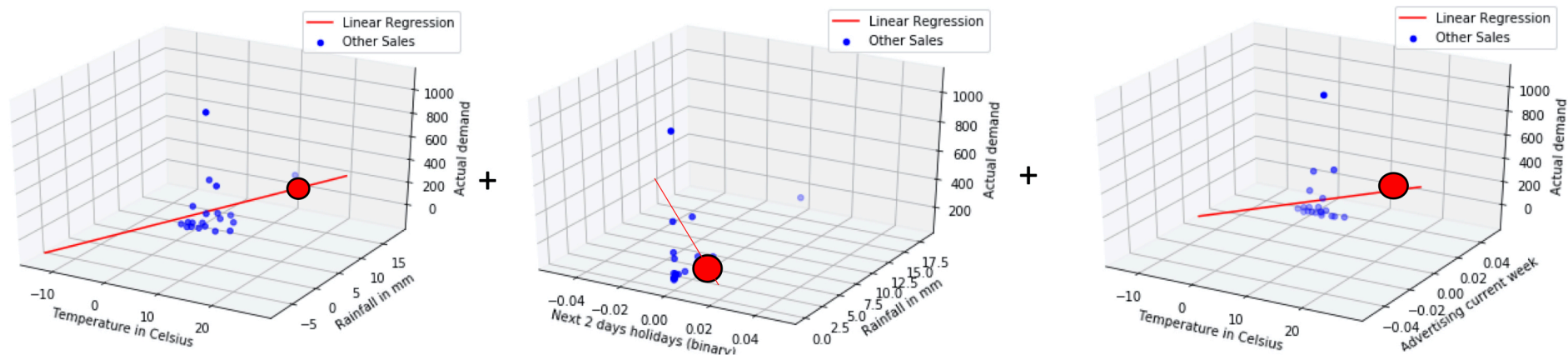
Linear regression is a statistical procedure that attempts to explain an observed **dependent variable** (actual demand) by **one or more independent variables/features** (e.g., temperature, holidays,...).

In linear regression, a linear model is assumed. Thus, only those correlations are used, in which the dependent variable is a **linear combination** of the regression coefficients (but not necessarily of the independent variables). The algorithm tries to **minimize the deviation (error term)** of the known cases from the calculated "regression line".

In the **example** on the left, according to the features *holidays* and *temperature* the model predicts the actual demand of buns for the next day.

Result: Bun Order Quantity

In order to show the explanation for the decision, multiple pairwise comparisons of two independent variables according to the dependent one is shown below. This is done for every possible variable's combination.



... for every Feature combination
=
Result: 228

Rate the following statement: "The presented explanations are good."

Please consider both, your perceived explanatory goodness of the algorithm's decision logic and its result.

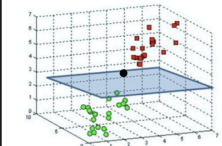
Strongly disagree Disagree More or less disagree Undecided More or less agree Agree Strongly Agree

<<

>>

67%

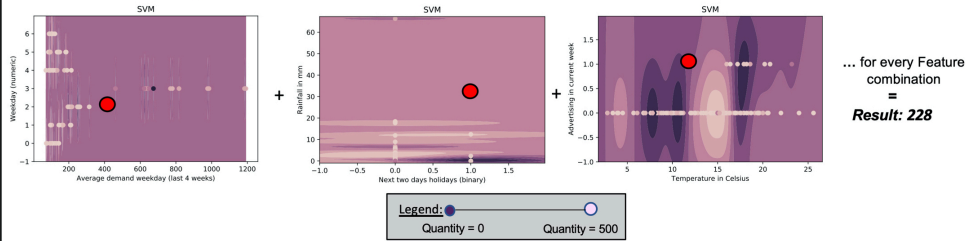
General Introduction: Algorithm of Support Vector Machines (SVM)



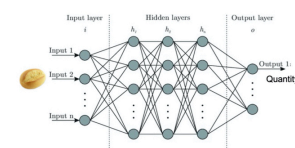
The SVM algorithm receives **data with n-features**, which are **represented as a n-dimensional vector**, whereby the individual features are assigned to a class. Based on the vector space, the algorithm determines the **optimum dividing hyperplanes (blue) between the features**, where the distance (margin) to the nearest points of the other class should be maximized. This wide, empty border later ensure that even objects that do not correspond exactly to the training objects are classified as reliably as possible. The hyperplanes classify the sales of buns according to its various features. When a particular observation is inserted into the SVM, the SVM predicts the sales of buns for tomorrow.

Result: Bun Order Quantity

In order to show the explanation for the decision, multiple pairwise comparisons of two features according to the bun order quantity is shown below. This is done for every possible feature combination.



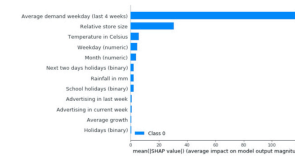
General Introduction: Algorithm of Artificial Neural Networks (ANN)



An artificial neural network (ANN) is an information processing paradigm that is inspired by the way the **biological nervous system** such as a human brain processes information. The algorithm is composed of a large number of highly **interconnected processing elements** (neurons) working in unison to solve a specific problem. The input layer receives new cases, which are calculated, through different mathematical operations, within the hidden layers. In the **example** on the left side sales data is forwarded to the ANN. Using many calculations in different layers, the model estimate the upcoming sales for buns.

Result: Bun Order Quantity

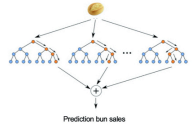
Feature importance of all observations:



On the left side, the mean global influence of the different features within the dataset in percent is shown. On the right side, the features with the influence (positive: red, negative: blue) on a specific prediction (model output value) based on the mean of all predictions (base value) in the dataset are presented.



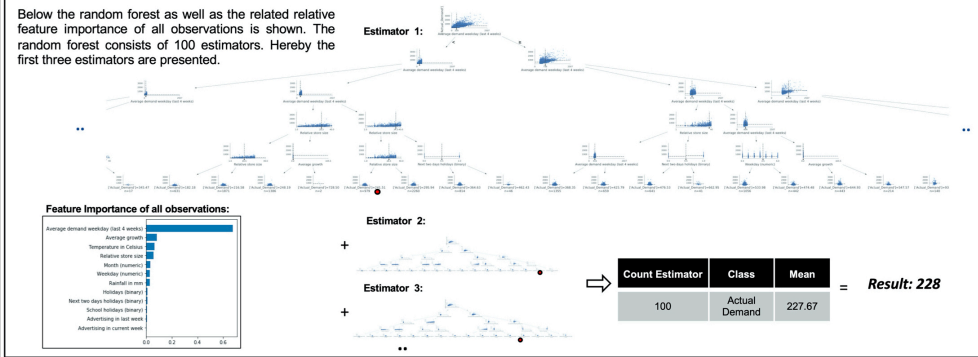
General Introduction: Algorithm of Random Forest



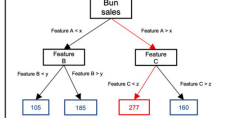
A random forest is a method consisting of **several uncorrelated decision trees also called estimators**. A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**. The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases. All decision trees have grown under a certain type of randomization during the learning process. For a prediction, each tree in that forest makes a decision, according to the features, and the **mean of all decision** determines the final result. In the **example** on the left side, the output would be the prediction of bun sales for tomorrow.

Result: Bun Order Quantity

Below the random forest as well as the related relative feature importance of all observations is shown. The random forest consists of 100 estimators. Hereby the first three estimators are presented.



General Introduction: Algorithm of Decision Trees



A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**.

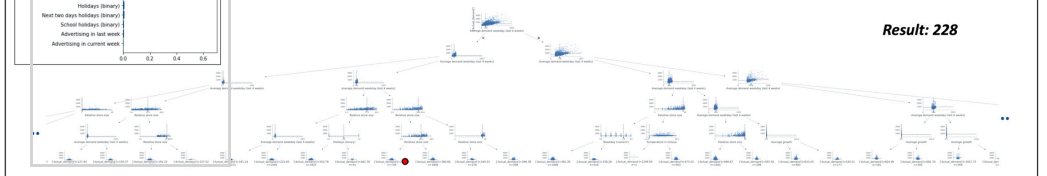
The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases. In the **example** on the left side, the output of the sales data would be 277 according to feature A and C (marked red).

Result: Bun Order Quantity

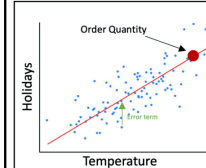
Feature importance of all observations:



The decision tree as well as the related relative feature importance of all observations is shown. The features of the inserted data are used to predict the result.



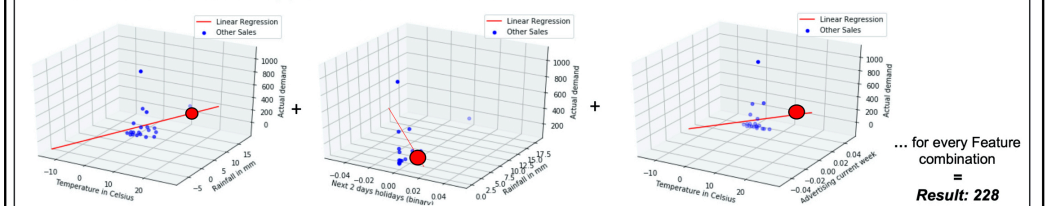
General Introduction: Algorithm of Linear Regression



Linear regression is a statistical procedure that attempts to explain an observed **dependent variable** (actual demand) by **one or more independent variables/features** (e.g., temperature, holidays,...). In linear regression, a linear model is assumed. Thus, only those correlations are used, in which the dependent variable is a **linear combination** of the regression coefficients (but not necessarily of the independent variables). The algorithm tries to **minimize the deviation (error term)** of the known cases from the calculated "regression line". In the **example** on the left, according to the features **holidays** and **temperature** the model predicts the actual demand of buns for the next day.

Result: Bun Order Quantity

In order to show the explanation for the decision, multiple pairwise comparisons of two independent variables according to the dependent one is shown below. This is done for every possible variable's combination.



Overall, which model with its related explanation would you choose as AI-based support system?

Choose

<<

>>



72%

Assignment

In the following, you will have to solve tasks on the comprehensibility of the presented ML algorithms and their predictions according to the shown use case. The structure of this assignment is shown below.

Use Case



Use Case: As reminder, the use case is shown again.

Algorithm & Prediction



Algorithm and Prediction: The heart of the ML model and the decision of it – as shown before.

Assignment

Task Type 1: Single Choice Task

Does A have a high influence in comparison to B?

☐

Choice 1

☐

Choice 2

☐

Choice 3

Task Type 2: Open Task

How does A relate to C, D and E?

Task Type 3: Cloze Task

In this sentence __ words are _____.

First missing word:

Second missing word:



Assignment: In order to check, if the AI-based system can help to solve tasks, we kindly ask you to answer 6 questions in total.

<<

>>



Lukas-Valentin
Herm



Jonas
Wanner



Prof. Dr.
Christian Janiesch



Franz Seubert



Prof. Dr.
Axel Winkelmann

Intro Assignment:

On the basis of your decision, your supervisor is requesting you to explain the models decision to him. To simplify this, the case as well as the model and the related explanation of your choice are shown once again.

Please read both of them again carefully. On the following page you will be interviewed on this subject. The tasks are used to satisfy their supervisor.



Imagine you are a baker and have to deal with the following problem:
Every day you have to decide **how many buns you want to order** for the sale on the next day. In order to make the task easier, **AI-based models** should be introduced in your company. To help your supervisor making the right choice, he or she will ask for your feedback on the various AI models.

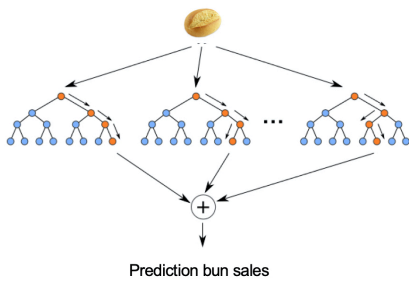
Each model was provided with the *delivery, returns and sales features as well as weather and calendar data of the last three years*.

Weekday (numeric)	Month (numeric)	School holidays (binary)	Holidays (binary)	Next 2 days holidays (binary)	Rainfall in mm	Temperature in Celsius	Average demand weekday (last 4 weeks)	Average growth (comparison last week)	Relative Store size	Advertising in current week (binary)	Advertising in last week binary	Actual Demand
1	1	1	1	0	11.02	31.20	34	1.23	0.52	1	0	156.2
2	2	0	1	0	10.56	34.00	234	1.51	0.56	0	0	185.9
3	7	1	0	1	11.00	34.89	345	1.53	0.62	1	0	236.3

Through this data, the various models now generate **order proposals** for the next day. The **explanations** of the models show how the *different features* influence the **sales quantity** of the buns.

Goal: Determine the optimal order quantity.
Setting: medium-stake, costs for wrong decision

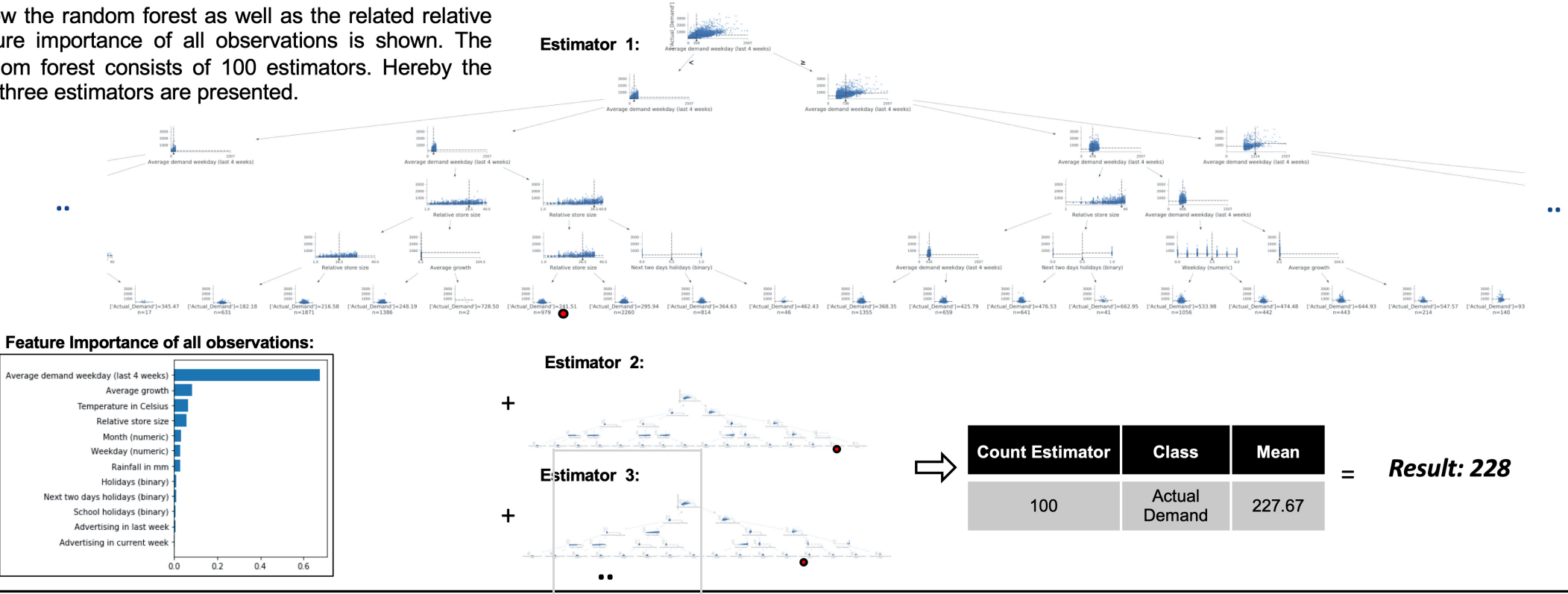
General Introduction: Algorithm of Random Forest



A random forest is a method consisting of **several uncorrelated decision trees also called estimators**. A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**. The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases. All decision trees have grown under a certain type of randomization during the learning process.
For a prediction, each tree in that forest makes a decision, according to the features, and the **mean of all decision** determines the final result. In the **example** on the left side, the output would be the prediction of bun sales for tomorrow.

Result: Bun Order Quantity

Below the random forest as well as the related relative feature importance of all observations is shown. The random forest consists of 100 estimators. Hereby the first three estimators are presented.



Does the average demand on a weekday (the last 4 weeks) have an influence on the quantity of buns being sold in this prediction?

- ☐ Yes
- ☐ No
- ☐ No information provided

The temperature in Celsius is in general more important than the rainfall in mm in the prediction.

- ☐ Yes
- ☐ No
- ☐ No information provided

Assignment Task (2/3)



Imagine you are a baker and have to deal with the following problem:
Every day you have to decide **how many buns you want to order** for the sale on the next day. In order to make the task easier, **AI-based models** should be introduced in your company. To help your supervisor making the right choice, he or she will ask for your feedback on the various AI models.

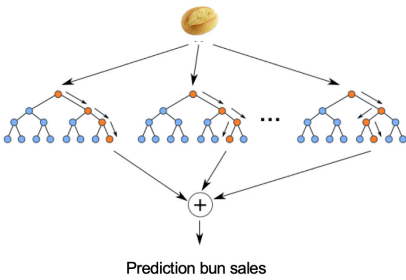
Each model was provided with the *delivery, returns and sales features as well as weather and calendar data of the last three years*.

Weekday (Jnumeric)	Month (numeric)	School holidays (binary)	Holidays (binary)	Next 2 days holidays (binary)	Rainfall in mm	Temperature in Celsius	Average demand weekday (last 4 weeks)	Average growth (comparison last week)	Relative Store size	Advertising in current week (binary)	Advertising in last week binary	Actual Demand
1	1	1	1	0	11.02	31.20	34	1.23	0.52	1	0	156.2
2	2	0	1	0	10.56	34.00	234	1.51	0.56	0	0	185.9
3	7	1	0	1	11.00	34.89	345	1.53	0.62	1	0	236.3

Through this data, the various models now generate **order proposals** for the next day. The **explanations** of the models show how the *different features* influence the **sales quantity** of the buns.

Goal: Determine the optimal order quantity.
Setting: medium-stake, costs for wrong decision

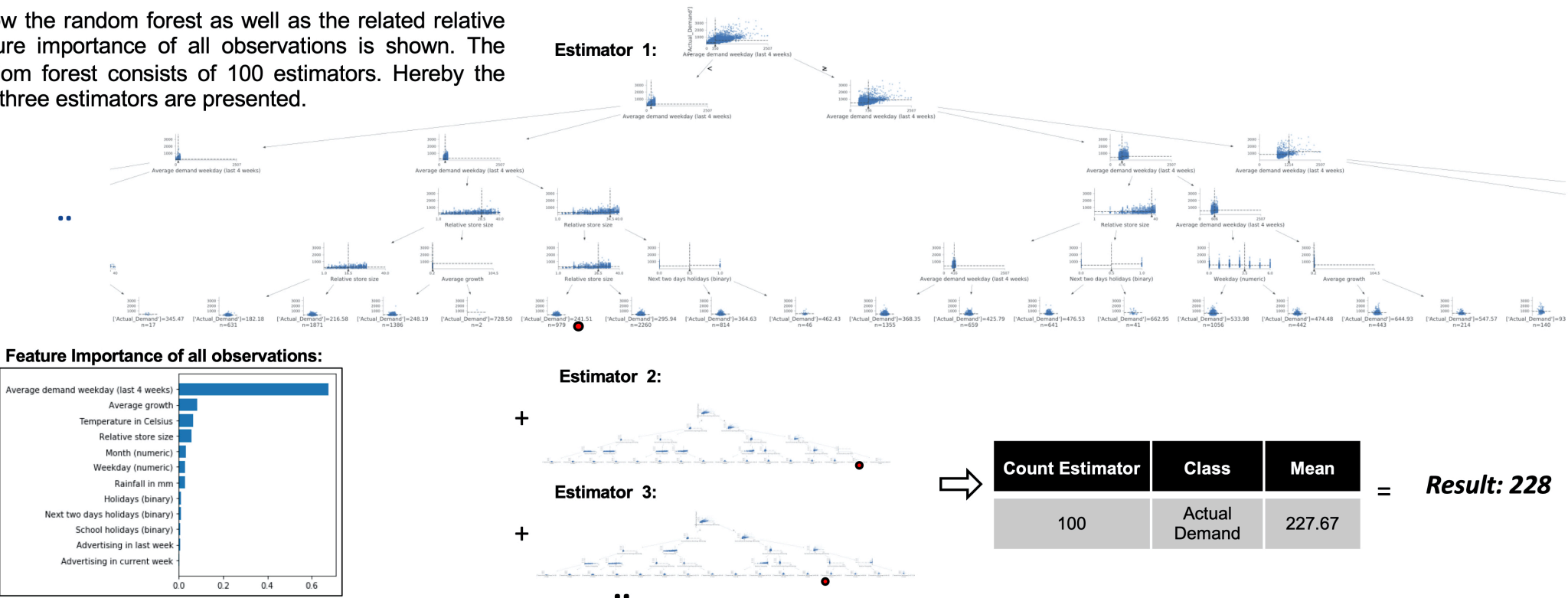
General Introduction: Algorithm of Random Forest



A random forest is a method consisting of **several uncorrelated decision trees also called estimators**. A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**. The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases. All decision trees have grown under a certain type of randomization during the learning process.
For a prediction, each tree in that forest makes a decision, according to the features, and the **mean of all decision** determines the final result. In the **example** on the left side, the output would be the prediction of bun sales for tomorrow.

Result: Bun Order Quantity

Below the random forest as well as the related relative feature importance of all observations is shown. The random forest consists of 100 estimators. Hereby the first three estimators are presented.



How does the overall influence of the average demand of the last 4 weeks differ from the specific prediction?

Your business partner, the local butcher, cannot estimate the needed quantity for the sale of steaks correctly. What features should the butcher take into account when planning future demand, assuming that buns and steaks have similar selling trends?



89%

Assignment Task (3/3)

The following questions should be completed without assistance. Therefore, please do not click on back.

Fill in the missing words of the following sentence below.

The explanations of the models show how the different _____ influence the _____ of buns.

First missing word:

Second missing word:

<<

>>



Lukas-Valentin
Herm



Jonas
Wanner



Prof. Dr.
Christian Janiesch



Franz Seubert

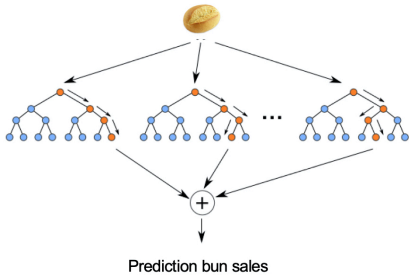


Prof. Dr.
Axel Winkelmann

Retrospective evaluation:

In the following, you will see your personal favorite once again for the sake of recollection.

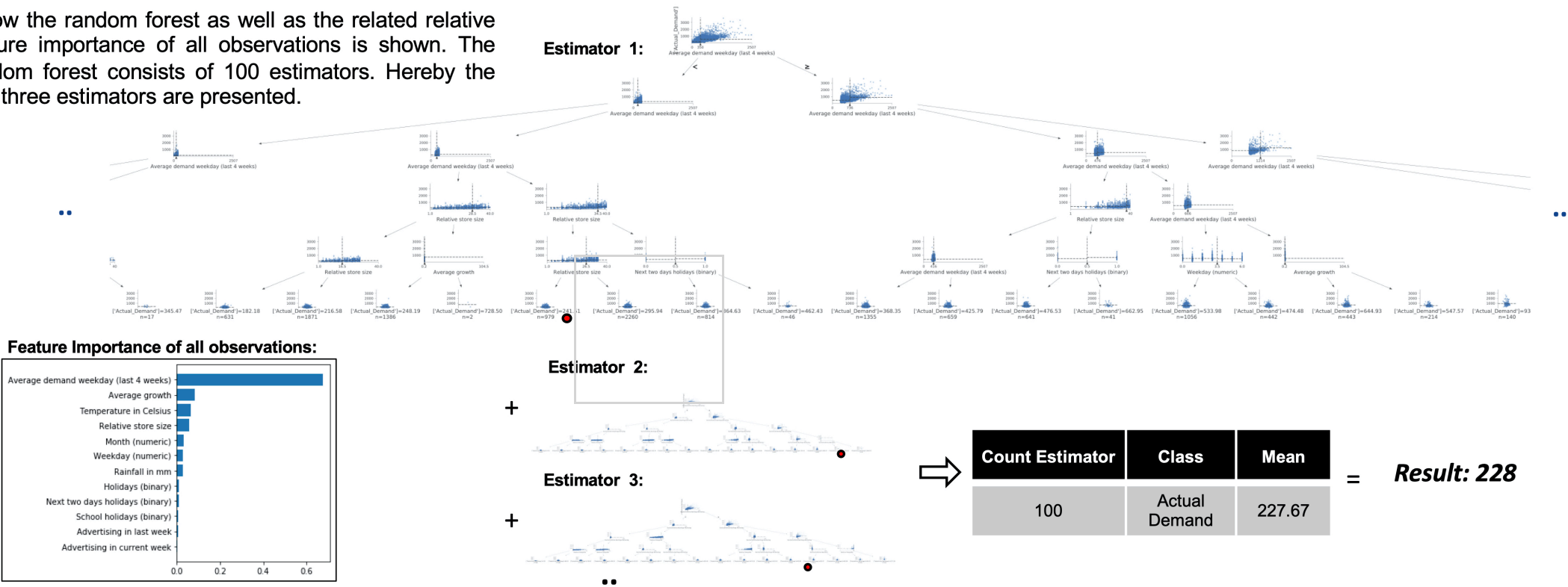
General Introduction: Algorithm of Random Forest



A random forest is a method consisting of **several uncorrelated decision trees also called estimators**. A decision tree represents a tree-like model of decisions and their possible consequences. It is an algorithm that only contains **conditional control statements**. The algorithm calculates a hierarchical decision construct of its conditional control statements to **minimize** the amount of **incorrect classified** training cases. All decision trees have grown under a certain type of randomization during the learning process. For a prediction, each tree in that forest makes a decision, according to the features, and the **mean of all decision** determines the final result. In the **example** on the left side, the output would be the prediction of bun sales for tomorrow.

Result: Bun Order Quantity

Below the random forest as well as the related relative feature importance of all observations is shown. The random forest consists of 100 estimators. Hereby the first three estimators are presented.



After you have answered the questions beforehand and thought critically about your selection:

What do you miss within the explanations, in order to feel confident in using this model as support system?



100%

Thanks for your participation.

Please find your referral link to prolific.co here: 



Lukas-Valentin
Herm



Jonas
Wanner



Prof. Dr.
Christian Janiesch



Franz Seubert



Prof. Dr.
Axel Winkelmann