

# Dynamic Pricing in der Luftfahrt

Analyse von Machine-Learning-Algorithmen zur Vorhersage von Flugpreisen

Kolloquium zur Bachelorarbeit

---

Dominik Lewin

Fernuniversität in Hagen

29.08.2023

- 1 Motivation
- 2 Grundlagen
- 3 Machine-Learning-Algorithmen
- 4 Praktischer Teil
- 5 Fazit und Ausblick

# Motivation

Motivation



Grundlagen



Machine-Learning-Algorithmen



Praktischer Teil



Fazit und Ausblick



Literatur



Ab Hamburg  
London

Mi., Aug. 16

**Ryanair**

125 €

Nur Hinreise · Economy Class · 1



16 Aug. · 14:45–15:15  
Ryanair · FR 1517

1 h 30 Min.  
HAM–STN

Nonstop



Preisentwicklung



[Buchungsoptionen ansehen](#)

Abbildung 1: Google Flüge: Preisentwicklung für einen Ryanair-Flug. Quelle: <https://www.google.com/travel/flights/>

Wie lässt sich Maschine Learning einsetzen,

- ▶ ... um Flugpreise vorherzusagen?

Wie lässt sich Maschine Learning einsetzen,

- ▶ ... um Flugpreise vorherzusagen?
- ▶ ... um eine Buchungsempfehlung (buchen / nicht buchen) abzugeben?

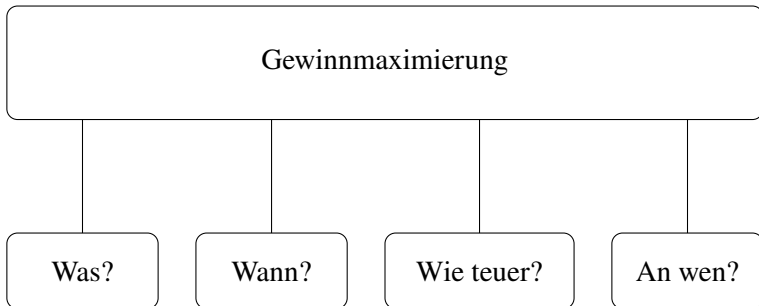


Abbildung 2: Grundgedanke des Revenue Managements. Eigene Darstellung.

## Wichtigste Anwendungsmerkmale:

- ▶ Weitgehend fixe Kapazitäten
- ▶ Verderblichkeit
- ▶ Nachfrageschwankungen

## Wichtigste Anwendungsmerkmale:

- ▶ Weitgehend fixe Kapazitäten
- ▶ Verderblichkeit
- ▶ Nachfrageschwankungen

## Instrumente:

- ▶ Preisdifferenzierung
- ▶ Kapazitätssteuerung
- ▶ Überbuchungssteuerung



# Preisdifferenzierung

Motivation  
○○

Grundlagen  
○○●○

Machine-Learning-Algorithmen  
oooooooooooo

Praktischer Teil  
ooooo

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Unterschiedliche Preise für gleiche Leistung
- ▶ Statische Preise bei herkömmlicher Preisdifferenzierung
- ▶ Festgelegter Preis im gesamten Verkaufszeitraum
- ▶ Eine dynamische Anpassung wäre sinnvoll

## Anwendungsmerkmale:

- ▶ Keine festgelegten Preise
- ▶ Geringer Aufwand bei Preisänderungen

## Anwendungsmerkmale:

- ▶ Keine festgelegten Preise
- ▶ Geringer Aufwand bei Preisänderungen

## Vorteile:

- ▶ Anpassung in Echtzeit
- ▶ Reaktion auf verändernde Zahlungsbereitschaft
- ▶ Reaktion auf Konkurrenzpreise
- ▶ Erlösmaximierung

# Machine-Learning-Algorithmen

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
●○○○○○○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Lineare Regression (R)
- ▶ Naiver Bayes-Klassifikator (K)
- ▶ Entscheidungsbäume (R,K)
- ▶ Random Forest (R,K)
- ▶  $k$ -Nächste Nachbarn (R,K)
- ▶ Support Vector Machine (R,K)
- ▶ Multilayer Perceptron (R,K)

R: Geeignet für  
Regressionsprobleme

K: Geeignet für  
Klassifikationsprobleme

# Lineare Regression

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
●○○○○○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

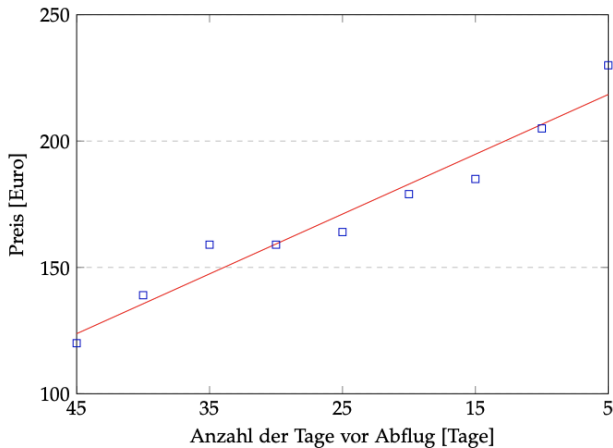


Abbildung 3: Fiktive Flugpreise in Abhängigkeit der Tage vor Abflug. Eigene Darstellung.

# Naiver Bayes-Klassifikator

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○●○○○○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Basiert auf dem Satz von Bayes aus der Wahrscheinlichkeitsrechnung
- ▶ Naive Annahme: Attribute sind unabhängig voneinander
- ▶ Beispiel-Flug: Ryanair, 150 Euro, 90 Tage vor Abflug
- ▶ Eingabevektor  $x = (\text{Ryanair}, 150, 90)$

$$P(\text{buchen}|x) = 0,81 \text{ und } P(\text{nicht buchen}|x) = 0,19 \rightarrow \text{buchen!} \quad (1)$$

# Entscheidungsbäume

Motivation  
○○

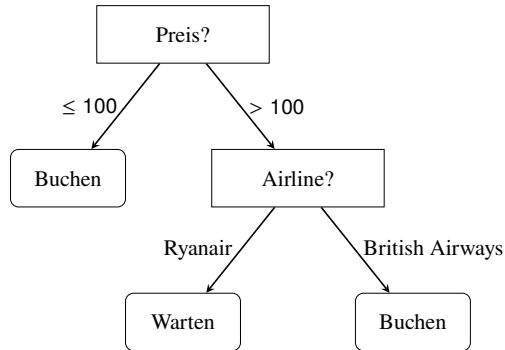
Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○●○○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○



- Entscheidungsbaum ist ein azyklischer Graph
- Der Graph lässt sich aus Trainingsdaten erzeugen
- Spaltkriterium zur Aufteilung benötigt

Abbildung 4: Entscheidungsbaum für ein fiktives Beispiel. Eigene Darstellung.

# Random Forest

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○●○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Wald aus vielen Entscheidungsbäumen
- ▶ Training auf unterschiedlichen Daten
- ▶ Beispiel-Flug: Ryanair, 150 Euro, 90 Tage vor Abflug



# Random Forest

Motivation  
○○

Grundlagen  
○○○○

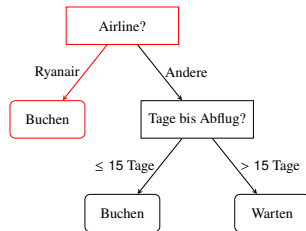
Machine-Learning-Algorithmen  
○○○○●○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Wald aus vielen Entscheidungsbäumen
- ▶ Training auf unterschiedlichen Daten
- ▶ Beispiel-Flug: Ryanair, 150 Euro, 90 Tage vor Abflug



Baum 1

# Random Forest

Motivation  
○○

Grundlagen  
○○○○

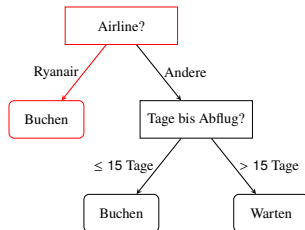
Machine-Learning-Algorithmen  
○○○○●○○○○

Praktischer Teil  
○○○○○

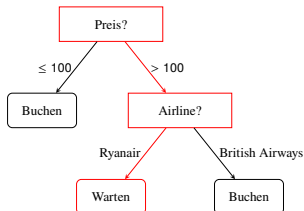
Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Wald aus vielen Entscheidungsbäumen
- ▶ Training auf unterschiedlichen Daten
- ▶ Beispiel-Flug: Ryanair, 150 Euro, 90 Tage vor Abflug



Baum 1



Baum 2

# Random Forest

Motivation  
○○

Grundlagen  
○○○○

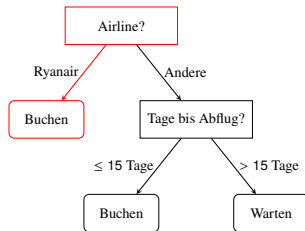
Machine-Learning-Algorithmen  
○○○○●○○○○

Praktischer Teil  
○○○○○

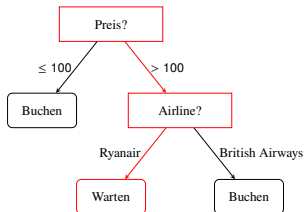
Fazit und Ausblick  
○○

Literatur  
○○

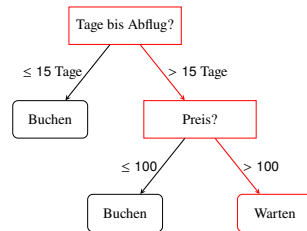
- ▶ Wald aus vielen Entscheidungsbäumen
- ▶ Training auf unterschiedlichen Daten
- ▶ Beispiel-Flug: Ryanair, 150 Euro, 90 Tage vor Abflug



Baum 1



Baum 2



Baum 3

Abbildung 5: Fiktiver Random Forest aus drei Entscheidungsbäumen. Eigene Darstellung.

# $k$ -Nächste Nachbarn

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○●○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- Annahme: Ähnliche Punkte liegen nah beieinander
- $k$  bezeichnet die Anzahl der Nachbarn
- Bestimmung der Distanz durch geeignete Abstandsmaße (z.B. Manhattan-Distanz)

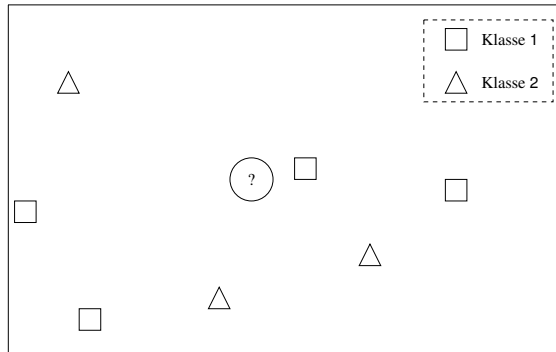


Abbildung 6: Klassifikation mit  $k = 1$  und  $k = 3$  Nachbarn. In Anlehnung an [Sri21].

# $k$ -Nächste Nachbarn

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○●○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Annahme: Ähnliche Punkte liegen nah beieinander
- ▶  $k$  bezeichnet die Anzahl der Nachbarn
- ▶ Bestimmung der Distanz durch geeignete Abstandsmaße (z.B. Manhattan-Distanz)

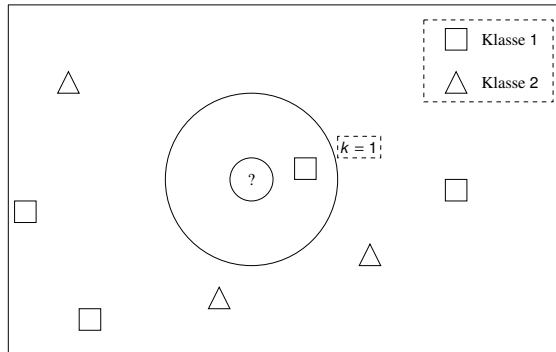


Abbildung 6: Klassifikation mit  $k = 1$  und  $k = 3$  Nachbarn. In Anlehnung an [Sri21].

# $k$ -Nächste Nachbarn

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○●○○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Annahme: Ähnliche Punkte liegen nah beieinander
- ▶  $k$  bezeichnet die Anzahl der Nachbarn
- ▶ Bestimmung der Distanz durch geeignete Abstandsmaße (z.B. Manhattan-Distanz)

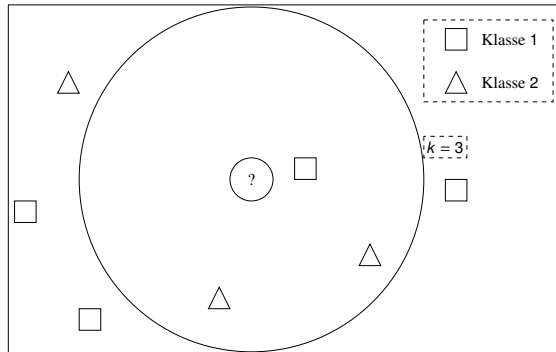


Abbildung 6: Klassifikation mit  $k = 1$  und  $k = 3$  Nachbarn. In Anlehnung an [Sri21].

# Support Vector Machine (SVM)

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○●○○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

- ▶ Hyperebene trennt die Klassen
- ▶ Maximaler Abstand zwischen Hyperebene und Datenpunkten
- ▶ *Stützvektoren (Support Vectors)* liegen auf dem Rand

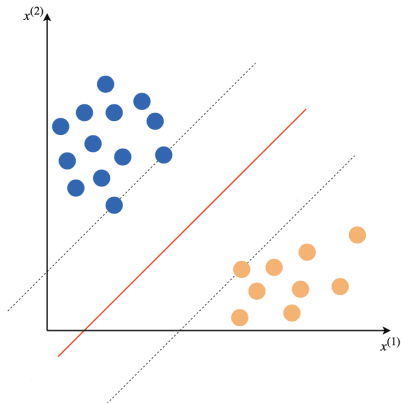


Abbildung 7: Zweidimensionales SVM-Modell. Entnommern aus [Bur19].

# Multilayer Perceptron (MLP)

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○○●○○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

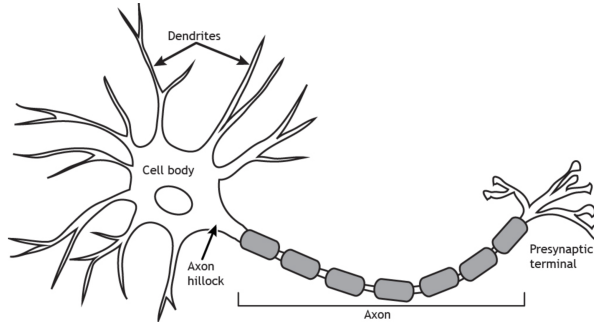


Abbildung 8: Biologisches Neuron. Entnommen aus [Hen21]



# Multilayer Perceptron (MLP)

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○○○●○

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

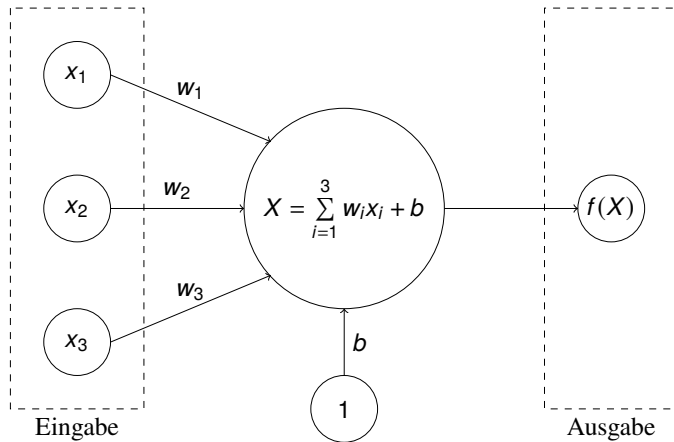


Abbildung 9: Perzeptron mit drei Eingaben. Eigene Darstellung.

# Multilayer Perceptron (MLP)

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○○○●

Praktischer Teil  
○○○○○

Fazit und Ausblick  
○○

Literatur  
○○

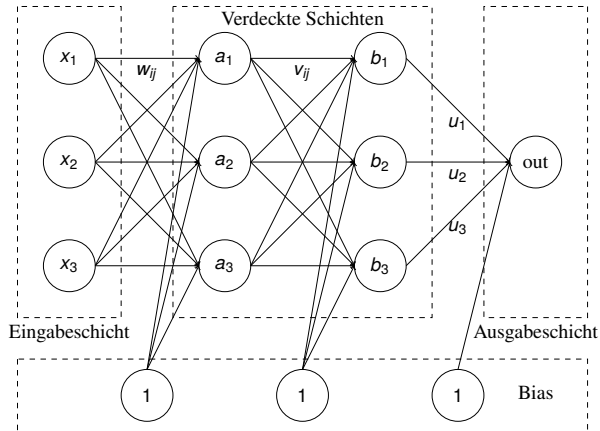


Abbildung 10: Mehrlagige Perzeptronen mit zwei verdeckten Schichten. Eigene Darstellung.

# Ziele des praktischen Teils

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○○○○○

Praktischer Teil  
●○○○○

Fazit und Ausblick  
○○

Literatur  
○○

EW 7462 (Hamburg - London), Abflug:  
29.10.2023, heutiger Preis: 160 Euro

# Ziele des praktischen Teils

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○○○○○

Praktischer Teil  
●○○○○

Fazit und Ausblick  
○○

Literatur  
○○

EW 7462 (Hamburg - London), Abflug:  
29.10.2023, heutiger Preis: 160 Euro

- ▶ Buchen?
- ▶ Warten?
- ▶ Ist das ein günstiger Preis?
- ▶ Wird der Preis noch günstiger?
- ▶ Wird der Preis morgen teurer?

Buchungsdatum	Abflugdatum	Abflughafen	Zielflughafen	Abflugzeit	Ankunftszeit	Preis	Flugnummer	Airline
2022-11-09	2023-02-26	HAM	LHR	17:55:00	18:40:00	98.0	EW 7460	Eurowings
2023-01-29	2023-02-26	HAM	LHR	13:20:00	14:05:00	120.0	EW 7462	Eurowings
2022-07-22	2023-02-27	HAM	LGW	13:05:00	13:45:00	35.0	U2 8342	easyJet
2023-02-23	2023-02-27	HAM	LHR	12:20:00	13:05:00	139.0	EW 7462	Eurowings
2022-09-08	2023-03-01	HAM	STN	21:20:00	21:55:00	32.0	FR 1519	Ryanair
2023-01-20	2023-03-02	HAM	LHR	12:20:00	13:05:00	80.0	EW 7462	Eurowings
2022-08-13	2023-03-03	HAM	STN	20:55:00	21:30:00	23.0	FR 1519	Ryanair
2022-12-27	2023-03-04	HAM	LHR	15:25:00	16:10:00	190.0	BA 975	British Airways

Abbildung 11: Auszug der Daten. Eigene Darstellung.

Lagemaß	Preis in Euro
Arithm. Mittel	77,53
Modalwert	58,00
Minimum	17,00
75 %- Quantil	85,00
Maximum	927,00

Tabelle 1: Darstellung ausgewählter Lagemaße. Eigene Darstellung

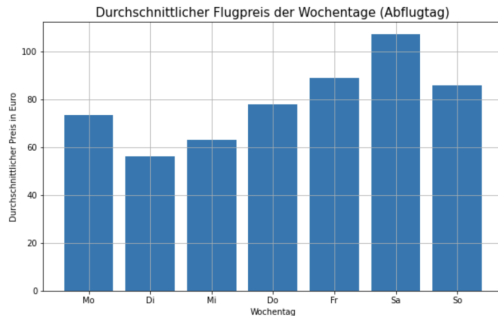


Abbildung 12: Durchschnittliche Flugpreise. Eigene Darstellung.

- ▶ Trainingsdaten 80%, Testdaten 20%
- ▶ Verwendet wurde scikit-learn (Python-Bibliothek)
- ▶ Training mit Standard-Parametereinstellungen
- ▶ Optimierung durch Anpassung der Parameter

# Evaluation

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
○○○○○○○○○○

Praktischer Teil  
○○○○●

Fazit und Ausblick  
○○

Literatur  
○○

	Lineare Regression	naiver Bayes	Entscheidungsbaum	Random Forest	k-Nearest Neighbors	SVM	MLP
<b>Regressionsproblem</b>							
Trainingsdaten-Score in %	(32,29)	——	64,32	66,53	66,78	(17,27)	64,92
Testdaten-Score in %	——	——	57,74	59,58	55,55	——	61,75
<b>Klassifikationsproblem</b>							
Trainingsdaten-Score in %	——	(66,49)	90,83	91,71	86,90	(68,70)	(84,15)
Testdaten-Score in %	——	——	90,67	91,66	87,87	——	——

Tabelle 2: Ergebnisse aller Algorithmen. Eigene Darstellung.



Wie lässt sich Maschine Learning einsetzen,

- ▶ ... um Flugpreise vorherzusagen?
- ▶ ... um eine Buchungsempfehlung (buchen / nicht buchen) abzugeben?

Wie lässt sich Maschine Learning einsetzen,

- ▶ ... um Flugpreise vorherzusagen?
- ▶ ... um eine Buchungsempfehlung (buchen / nicht buchen) abzugeben?

Ausblick für zukünftige Arbeiten:

- ▶ Deutlich größerer Datensatz
- ▶ Airlinespezifische Trainings
- ▶ Parameter optimieren
- ▶ Zusätzliche Attribute

# Zusammenfassung

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
oooooooooooo

Praktischer Teil  
ooooo

Fazit und Ausblick  
○●

Literatur  
○○

- ▶ Revenue Management dient der Gewinnmaximierung
- ▶ Preisdifferenzierung als Instrument des Revenue Managements
- ▶ Art der Preisdifferenzierung → Dynamic Pricing
- ▶ Praktischer Teil: Anwendung von 7 Machine-Learning-Algorithmen
- ▶ Schlechte Ergebnisse bei Flugpreisvorhersage (Multilayer Perceptron mit 61,75%)
- ▶ Gute Ergebnisse bei Buchungsempfehlung (buchen / nicht buchen) (Random Forest mit 91,66%)

# Zusammenfassung

Motivation  
○○

Grundlagen  
○○○○

Machine-Learning-Algorithmen  
oooooooooooo

Praktischer Teil  
ooooo

Fazit und Ausblick  
○●

Literatur  
○○

- ▶ Revenue Management dient der Gewinnmaximierung
- ▶ Preisdifferenzierung als Instrument des Revenue Managements
- ▶ Art der Preisdifferenzierung → Dynamic Pricing
- ▶ Praktischer Teil: Anwendung von 7 Machine-Learning-Algorithmen
- ▶ Schlechte Ergebnisse bei Flugpreisvorhersage (Multilayer Perceptron mit 61,75%)
- ▶ Gute Ergebnisse bei Buchungsempfehlung (buchen / nicht buchen) (Random Forest mit 91,66%)

Danke für die Aufmerksamkeit!  
Noch Fragen?

# Literaturverzeichnis

Motivation ○○	Grundlagen ○○○○	Machine-Learning-Algorithmen ○○○○○○○○○○	Praktischer Teil ○○○○○	Fazit und Ausblick ○○	Literatur ○○
[Bur19]	Andriy Burkov. <i>Machine Learning kompakt: Alles, was Sie wissen müssen</i> . MITP-Verlags GmbH & Co. KG, 2019.				
[Hen21]	Casey Henley. <i>Foundations of Neuroscience</i> . Michigan State University Libraries, 2021.				
[Pra22]	Naveen Prasath, Sherin Eliyas und Sathish Kumar. “A Prediction of Flight Fare Using K-Nearest Neighbors”. In: <i>2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)</i> . IEEE. 2022, S. 1347–1351.				
[Ren15]	Ruixuan Ren, Yunzhe Yang und Shenli Yuan. <i>Prediction of airline ticket price</i> . Techn. Ber. University of Stanford, 2015.				
[Sri21]	Abhishek Srivastava. “Impact of k-nearest neighbour on classification accuracy in knn algorithm using machine learning”. In: <i>Advances in Smart Communication and Imaging Systems: Select Proceedings of MedCom 2020</i> . Springer. 2021, S. 363–373.				
[Tzi17]	Konstantinos Tziridis u. a. “Airfare prices prediction using machine learning techniques”. In: <i>25th European Signal Processing Conference (EUSIPCO)</i> . IEEE. 2017, S. 1036–1039.				

LH 500 F8 A5 OL J9 C9 D9 Z9 /FRA 1 GIG 2 2215 0455+1E0/74H 11:40  
P9 IL G9 E9 NL RL Y9 B9 M9 U9 H9 XL Q9 V9 W9 S9 TC LC KC

Abbildung 13: Buchungsklassen eines Lufthansa-Fluges. Quelle: <https://meilenoptimieren.com/alles-ueber-buchungsklassen/>

Buchungsklassen der First Class: F, A, O

Buchungsklassen der Business Class: J, C, D, Z, P, I

Buchungsklassen der Premium Economy Class: G, E, N, R

Buchungsklassen der Economy Class: Y, B, M, U, H, X, Q, V, W, S, T, L, K

