



Certificate of Advanced Studies

Practical Machine Learning

Den intelligenten Software-Systemen gehört die Zukunft. Sie analysieren Daten und Situationen, erstellen Prognosen, treffen Entscheide oder geben Empfehlungen ab und steuern Prozesse. Das CAS Practical Machine Learning (CAS PML) rüstet Sie mit den Methoden und Werkzeugen aus, damit Sie solche Systeme bauen, einsetzen und nutzen können.



ti.bfh.ch/cas-pml

Inhaltsverzeichnis

1	Umfeld	3
2	Zielpublikum	3
3	Ausbildungsziele	3
4	Voraussetzungen	3
5	Termine, Anmeldung und Durchführungsort	4
6	Kompetenzprofil	4
7	Kursübersicht	5
8	Kursbeschreibungen	6
	8.1 R-Update	7
	8.2 Einführung in die Denk- und Handlungsweise des ML	7
	8.3 Phyton	8
	8.4 Feature Engineering	8
	8.5 Überwachtes Lernen	9
	8.6 Bewertung von Modellen	9
	8.7 Nicht-überwachtes Lernen	10
	8.8 Neuronale Netze	10
	8.9 Wahlfach Image Analysis	11
	8.10 Wahlfach Deep Learning	11
	8.11 Wahlfach Text Analytics und Natural Language Processing	12
	8.12 Wahlfach Social Network Analytics	12
	8.13 Wahlfach Recommender Systems	13
	8.14 Projektarbeit	13
9	Kompetenznachweis	15
10	Lehrmittel	15
11	Dozierende	17
12	Organisation	17

1 Umfeld

Die Anforderungen an Optimierung, Effizienz und Qualität von Unternehmensprozessen und Dienstleistungen wächst laufend. Mit Machine Learning (ML) verschaffen Sie sich Einblicke in die Daten Ihrer Organisation, optimieren Ihre Dienstleistungen und Unternehmensprozesse und können präzisere Entscheide treffen. Die Anwendungen von ML sind äußerst vielseitig, beispielsweise:

- Präzisierung von Verbrauchs- und Beschaffungsprognosen in der Logistik
- Automatische Klassifikation von Dokumenten im Kontext von e-Government
- Produktempfehlung in der Kundenberatung
- Erkennung von problematischen Teilprozessen in der medizinischen Versorgung
- Analyse und Steuerung von IoT-Netzwerken
- Vorhersage von Störfällen in der industriellen Produktion
- Überwachungsaufgaben in der IT-Sicherheit
- Erkennen und Klassifizieren von Objekten in Bildern oder Tonmustern in Audiodaten

Es stehen einfach zu bedienende und mächtige Softwarepakete für die Anwendung von Machine Learning zur Verfügung. Deren sinnvolle Nutzung erfordert allerdings Kenntnisse über die Wirkungsweise, die Rahmenbedingungen und die Qualitätsmessung der verschiedenen Algorithmen.

2 Zielpublikum

Das CAS PML richtet sich an:

- Informatikerinnen und Informatiker, die Machine Learning-Komponenten kompetent in Software-Anwendungen integrieren wollen.
- Fachexpertinnen, Fachexperten aus IT und Business, die Machine Learning Techniken kennenlernen und anwenden möchten.

3 Ausbildungsziele

Dieses CAS befähigt Sie zur Anwendung von ML im eigenen Arbeitsgebiet, sowie zur Mitarbeit in Teams, die ML-Methoden einsetzen oder implementieren:

- Sie kennen Einsatzbereiche und Anwendungsmöglichkeiten von Machine Learning.
- Sie können in der Evaluation, bei der Planung und im Einsatz von Machine Learning-Komponenten kompetent mitentscheiden.
- Sie können die Leistungsfähigkeit eines Machine Learning-Algorithmus bewerten.

4 Voraussetzungen

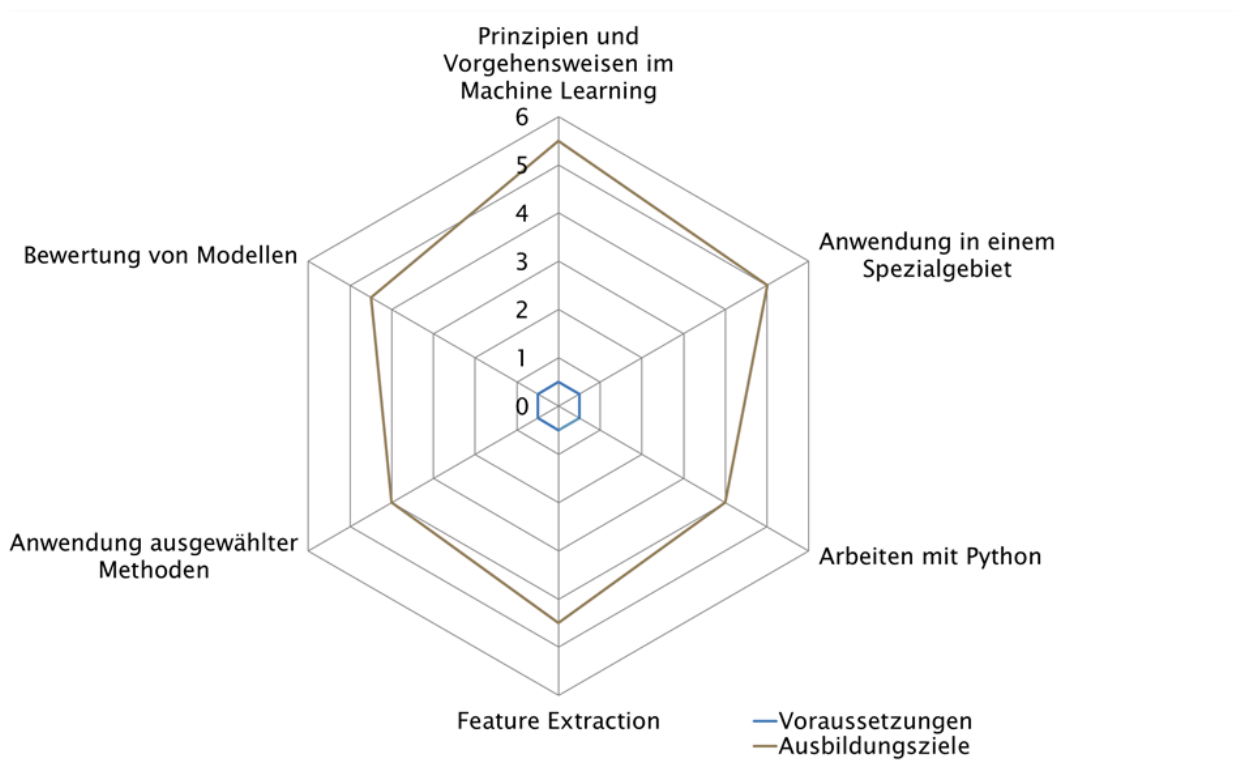
- Sie bringen ein Bachelorstudium mit, typischerweise in Informatik-, Ökonomie- oder Engineering-Disziplinen und haben Spass an einer algorithmischen Denkweise.
- Notwendig sind Vorkenntnisse in Statistik und Datenanalyse und der Sprache R (www.r-project.org), entsprechend etwa dem Stoff im CAS Datenanalyse.
- Für Software-Ingenieure und -Ingenieurinnen, die mit Skriptsprachen umfangreiche Erfahrung haben, oder für alle, die R schon kennen, aber lange nicht verwendet haben, wird ein 1 tägiger R-Update angeboten.

5 Termine, Anmeldung und Durchführungsort

Kursstart Kalenderwoche 43/2019, Anmeldeschluss Ende Kalenderwoche 40/2019
Das CAS dauert ein Semester und findet gemäss Stundenplan statt.

Berner Fachhochschule, Weiterbildung, Wankdorffeldstrasse 102, 3014 Bern,
Telefon +41 31 848 31 11, E-Mail office.ti-be@bfh.ch.

6 Kompetenzprofil



Kompetenzstufen

1. Kenntnisse/Wissen
2. Verstehen
3. Anwenden
4. Analyse
5. Synthese
6. Beurteilung

7 Kursübersicht

Kurs / Lehreinheit	Lektionen	Stunden	Dozierende
R-Update (optional)	8		Werner Dähler
Einführung in die Denk- und Handlungsweise des Machine Learning	16		Jürgen Vogel
Überwachtes Lernen	16		Werner Dähler
Bewertung von Modellen	8		Werner Dähler
Feature Engineering	8		Werner Dähler
Python	24		Niklaus Johner
Nicht-überwachtes Lernen	16		Matthias Dehmer
Neuronale Netze	8		Matthias Dehmer
Wahlfach: Image Analysis	24		Marcus Hudritsch
Wahlfach: Deep Learning	24		Romeo Kienzler
Wahlfach: Text Analytics und Natural Language Processing (NLP)	24		Jürgen Vogel
Wahlfach: Social Network Analytics	24		Debra Hevenstone
Wahlfach: Recommender Systems	24		Andreas von Ballmoos
Projektarbeit	8	90	Div. Betreuer
Total	152	90	

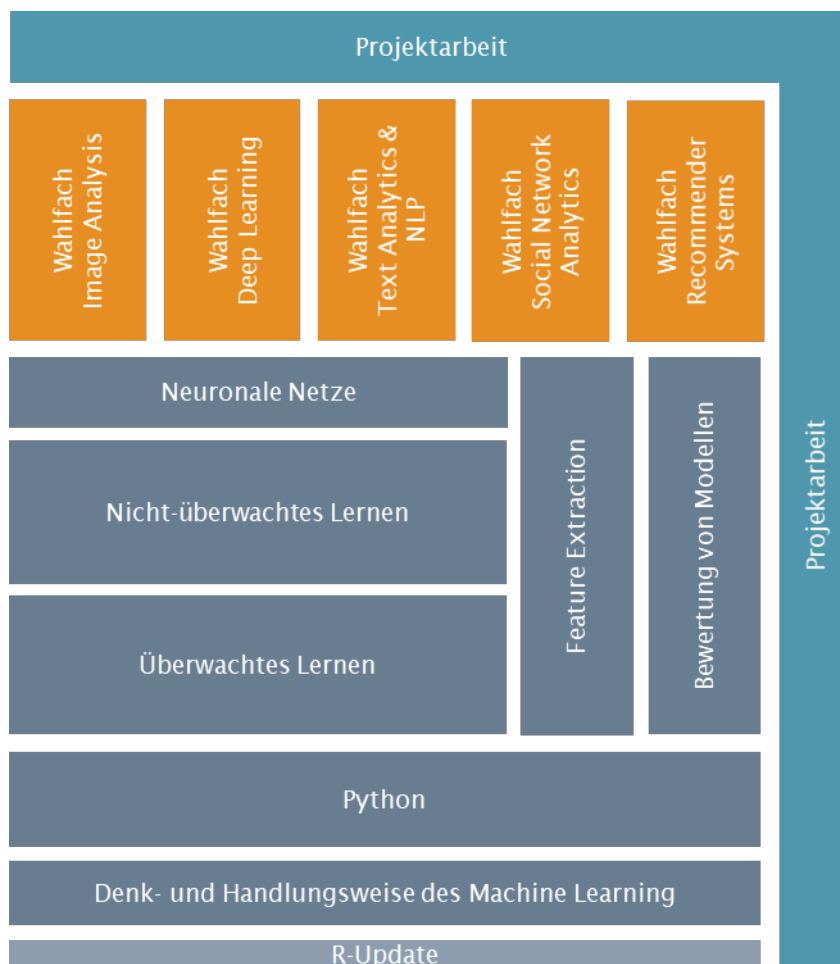
Das CAS umfasst insgesamt 12 ECTS-Punkte. Für die einzelnen Kurse ist entsprechend Zeit für Selbststudium, Prüfungsvorbereitung etc. einzurechnen.

8 Kursbeschreibungen

Das CAS beginnt mit einer generellen Einführung in die Denk- und Handlungsweise des Machine Learning, gefolgt von einem Kurs in Python, einem der zentralen Werkzeuge in diesem Fachgebiet. Anschliessend werden die beiden klassischen Lernmethoden, überwachtes und nicht-überwachtes Lernen behandelt. Feature Extraction ist eine entscheidende Tätigkeit für alle Arten von Algorithmen, um einen Datenbestand für die Analyse geeignet vorzubereiten. Die Bewertung von Modellen ermöglicht quantitative Aussagen, welche Lernmodelle und spezifischen Algorithmen am Besten geeignet sind.

Neuronale Netze haben einen grossen Aufschwung genommen für verschiedenste Problemlösungen, insbesondere in der Bildanalyse, für iterative Verfahren im Deep Learning und Reinforcement Lernen in der Robotik. In diesem Kurs wird der Grundstein für Anwendungen mit neuronalen Netzen gelegt. Die Wahlfächer ermöglichen eine Vertiefung in einem spezifischen Gebiet. Teilnehmerinnen und Teilnehmer besuchen ein oder mehrere Wahlfächer.

Eine Projektarbeit zu einem Thema nach Wahl, optimalerweise im eigenen beruflichen Umfeld, begleitet und vertieft die Lerninhalte.



Nachfolgend sind die einzelnen Kurse dieses Studienganges beschrieben.

Der Begriff Kurs schliesst alle Veranstaltungstypen ein, es ist ein zusammenfassender Begriff für verschiedene Veranstaltungstypen wie Vorlesung, Lehrveranstaltung, Fallstudie, Living Case, Fach, Studienreise, Semesterarbeiten usw.

8.1 R-Update

Kurzübersicht zur Verwendung von R und R-Studio. Dieser Kurs ersetzt nicht eine vollständige Einführung in die Sprache R. Angesprochen sind erfahrene Software-Ingenieure, die sich in einem Quick-Start die wichtigsten Sprachelemente von R aneignen wollen. Der Kurs richtet sich ebenfalls an alle, die R kennen, aber ev. schon lange nicht mehr verwendet haben. Für die weiteren Kurse des CAS werden R-Kenntnisse vorausgesetzt.

Lernziele	Die Teilnehmenden können mit R resp. R-Studio vorliegende Daten aus unterschiedlichen Quellen und in unterschiedlichen Datenformaten einlesen, sichten und für nachgeschaltete Analysen aufbereiten. Sie kennen die Grundzüge der R Syntax und sind in der Lage, selber einfache Funktionen zu schreiben.
Themen und Inhalte	<ul style="list-style-type: none">– Grundlagen der R Syntax– R Objekte: Variablen, Listen, Matrizen, Data Frames– Datenbezug aus unterschiedlichen Quellen– Basisoperationen– Datenmanipulationen, Selektion und Transformation– Grundlagen der Programmierung: Kontrollstrukturen, Funktionen– Paketverwaltung– Praktische Übungen in R
Lehrmittel	<ul style="list-style-type: none">– Folien/Skript– Abgegebene Beispiele in R Code– Literaturempfehlungen Nr. 4, 5

8.2 Einführung in die Denk- und Handlungsweise des ML

Lernziele	Die Teilnehmenden bekommen eine erste Orientierung zu Themen und Anwendungsbereichen von Machine Learning (ML). Sie lernen die typische Denk- und Vorgehensweise in ML-Projekten kennen und können entsprechend planen und handeln.
Themen und Inhalte	<ul style="list-style-type: none">– Was ist Machine Learning?– Anwendungsbereiche und Anwendungsmuster– Arten von ML-Algorithmen– Typische Vorgehensweise<ul style="list-style-type: none">– Fragestellung definieren– Daten beschaffen– Daten aufbereiten– Methodik und Algorithmen auswählen– Arbeiten mit Lerndaten und Testdaten– Resultate evaluieren– Arbeiten mit Standarddatensätzen– Plattformen der ML-Community– Erste praktische Übungen mit Python zu ausgewählten Beispieldaten– Einführung der Methoden Decision Trees, K-Means, Naives Bayes
Lehrmittel	<ul style="list-style-type: none">– Folien/Skript– Literaturempfehlungen Nr. 1, 2, 3

8.3 Phyton

Einführung in die Sprache mit Bezug auf Machine Learning Anwendungen. Die Einführung basiert auf Python 3.0.

Lernziele	Die Teilnehmenden lernen die Grundlagen der Python Programmierung Sprache. Sie können Python Skripte lesen, schreiben und ausführen. Sie kennen einen Teil der Standardbibliothek sowie die wichtigsten Bibliotheken für wissenschaftliches Rechnen. Sie können Dateien lesen und schreiben, sowie Daten bearbeiten und grafisch darstellen.
Themen und Inhalte	<ul style="list-style-type: none">– Datentypen und Strukturen– Operatoren und Vergleiche– Ablaufsteuerung (Schleifen, Bedingungen ...)– Zeichenketten (strings) Formatierung und Parsing– Dateien lesen und schreiben– Wissenschaftliches Rechnen mit numpy und scipy, scikit– Einführende Techniken des Feature Engineering.– Grafische Darstellung von Daten mit matplotlib– Praktische Übungen zu allen Themen– Beispiele zum Datendatein lesen, Daten bearbeiten und Funktion fitten.– Pandas
Lehrmittel	<ul style="list-style-type: none">– Folien/Skript– Literaturempfehlungen Nr. 9, 10, 11

8.4 Feature Engineering

Lernziele	Die Teilnehmenden erkennen die zentrale Stellung von Feature Selektion und Extraktion im Machine Learning Prozess, und können verschiedene Verfahren zur Aufbereitung ihrer Daten anwenden.
Themen und Inhalte	<ul style="list-style-type: none">– Warum gute Features im Machine Learning wichtig sind– Analyse der vorliegenden Features– Wahl geeigneter Attribute aus vorhandenen Datensätzen (Feature Selektion)– Feature Extraktion durch Transformation– Dimensionsreduktion– Übungen in R– Anwendungsbeispiele
Lehrmittel	<ul style="list-style-type: none">– Folien/Skript– Literaturempfehlungen Nr. 6, 7, 8

8.5 Überwachtes Lernen

Überwachtes Lernen ist eine klassische Disziplin des Machine Learning. Einem Algorithmus wird anhand von Trainingsdaten ein bestimmtes Verhalten beigebracht, welches er dann auf neue Daten anwendet.

Lernziele	Die Teilnehmenden kennen die Prinzipien und einige typische Algorithmen aus dem Bereich des überwachten Lernens. Sie können Algorithmen aufgrund einer Fragestellung als angemessen auswählen und anwenden.
Themen und Inhalte	<ul style="list-style-type: none"> – Anwendungen und Eigenschaften von häufig eingesetzten Methoden: <ul style="list-style-type: none"> – Einführung K Nearest Neighbours (KNN) – Einführung in Decision Trees – Einführung Logistische Regression – Einführung Random Forest – Gegenüberstellung dieser und weiterer Klassifikationsmethoden – Verschiedene Regressionsmethoden im Überblick – Praktische Übungen in R/Python <ul style="list-style-type: none"> – Beispiele aus den R-Paketen und an einem Fallstudien Dataset
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript – Literaturempfehlungen Nr. 6, 8

8.6 Bewertung von Modellen

Wie gut ist ein Machine Learning Algorithmus im Vergleich zu einem anderen? Wie stabil oder verlässlich ist ein Algorithmus, wenn man ihm unterschiedliche Daten zu verarbeiten gibt? Wie misst man die Güte eines Machine Learning Modells als Kombination von Algorithmus, Daten und Parametrisierung? Um diese Frage kümmert sich die Bewertung von Modellen (Modellevaluation).

Lernziele	Die Teilnehmenden lernen, dass die Güte und die Nützlichkeit eines Machine Learning Modelles mit bestimmten Kriterien gemessen und bewertet werden muss. Sie können je nach Situation ein geeignetes Verfahren anwenden und einige wichtige Kennzahlen verstehen und interpretieren.
Themen und Inhalte	<ul style="list-style-type: none"> – Datenreduktion mittels Sampling, Stratified Sampling, Reservoir Sampling – Verfahren und Kennzahlen für überwachtes Lernen (Klassifikation) – Kreuzvalidierung (Crossvalidation) und Bootstrapping – Precision, Recall, Accuracy, F-Score – ROC und AUC – Log Loss – Kennzahlen für Regression-Verfahren – Praktische Übungen in R – Beispiele aus den R-Paketen – Anwendungsbeispiele aus der Bioinformatik und der Medizin
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript – Literaturempfehlungen Nr. 6, 8

8.7 Nicht-überwachtes Lernen

Beim klassischen, nicht-überwachten Lernen versucht ein Algorithmus typischerweise, Daten in Cluster einzuteilen, oder Abhängigkeiten zwischen Datenwerten zu finden. Der Nutzer kennt in der Regel die Cluster nicht im Voraus, sondern will anhand der automatischen Analyse neue Erkenntnisse gewinnen. Beispielsweise können Kundensegmentierungen, Bildanalysen oder Empfehlungssysteme mit nicht-überwachtem Lernen realisiert werden.

Lernziele	Die Teilnehmenden kennen die Prinzipien und einige typische Algorithmen aus dem Bereich des überwachten Lernens. Sie kennen Einsatzmöglichkeiten und können einen Algorithmus aufgrund der Fragestellung wählen und anwenden.
Themen und Inhalte	<ul style="list-style-type: none"> – Partition Around Medoids (PAM), K-Means Clustering – Hierarchisches Clustering – Recommender Systeme und kollaboratives Filtern <ul style="list-style-type: none"> – Speicher- und Modellbasierte Verfahren – Matrix Faktorisierung, latente Faktoren – Explizite und implizite Bewertungen – Assoziationsanalyse <ul style="list-style-type: none"> – Warenkorbanalyse – Apriori Algorithmus – Anwendungsbeispiele <ul style="list-style-type: none"> – Detektion von Ausreißern (Anomaly Detection) – Film-Empfehlungen (ähnlich zu Netflix) anhand des MovieLens Datensatz – Kundensegmentierung – Übungen in Python mit scikit-learn
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript – Literaturempfehlungen Nr. 12, 13, 14, 15

8.8 Neuronale Netze

Die Idee der Neuronale Netze entspringt einem vereinfachten Modell des Aufbaus und der Lernvorgänge eines biologischen Gehirns. Mit den heutigen Rechenkapazitäten können Neuronale Netze ihr Potential immer besser ausschöpfen. Viele Erfolge in der Robotik, der Analyse von Bildern und dem dynamischen Lernen, machen Neuronale Netze zu einem populären Anwendungs- und Forschungsgebiet.

Lernziele	Die Teilnehmenden kennen die Prinzipien und einige typische Arten von Neuronalen Netzen. Sie können grob abschätzen, welche Netze für welche Aufgaben in Frage kommen. Dieser Kurs dient als Vorbereitung für die weitere Vertiefung und das Arbeiten mit neuronalen Netzen in den Wahlkursen, insbesondere Image Analysis und Deep Learning.
Themen und Inhalte	<ul style="list-style-type: none"> – Ursprung und Idee – Funktionsweise von Neuronalen Netzen – Arten und Typologien – Einsatzmöglichkeiten, Beispiele, Success Stories – Schwierigkeiten mit neuronalen Netzen
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript

8.9 Wahlfach Image Analysis

Hauptfokus dieses Kurses ist die Erkennung von Objekten und die fachliche Beschreibung von Bilddaten (Feature Extraction).

Lernziele	Die Teilnehmenden erhalten eine Übersicht, wie Bilder analysiert und danach klassischen Machine Learning Methoden zugeführt werden können.
Themen und Inhalte	<ul style="list-style-type: none"> – Introduction – Image Segmentation – Region Representation – Feature Extraction – Classification – Geometric Analysis
Lehrmittel	<ul style="list-style-type: none"> – Detailliertes Skript und Folien – Übungen in Matlab

8.10 Wahlfach Deep Learning

Deep Learning ist eines der Top IT Themen. Deep Learning ist klassischen Machine Learning Algorithmen in verschiedenen Anwendungsbereichen überlegen. Dazu gehören Bilderkennung, Textanalyse (NLP), Anomaly Detection, Zeitreihenvorhersagen und vieles mehr. In diesem Modul wollen wir verstehen wie künstliche neuronale Netze aufgebaut sind und wie diese trainiert und angewendet werden.

Lernziele	<p>Prinzipien und Anwendungen, Arbeiten mit neuronalen Netzen, Einsatz für Bilderkennung und mehr.</p> <p>Die Teilnehmenden:</p> <ul style="list-style-type: none"> – Können ein Neuronales Netz selbständig in Python implementieren und trainieren – Verstehen die nötigen Grundlagen in Linearer Algebra – Führen selbständig ein Neuronales Netzwerk Training durch – Verstehen die Anwendungsgebiete für DeepLearning – Können diverse DeepLearning Frameworks evaluieren und einsetzen
Themen und Inhalte	<ul style="list-style-type: none"> – Grundlagen in Linearer Algebra – Grundlagen von neuronalen Netzwerken. – Kostenfunktionen und Training mit Gradient Descent. – Einführung in TensorFlow, Keras und Apache SystemML. – Übungen – Selbststudium – Anwendungsbeispiele <ul style="list-style-type: none"> – Handschriftenerkennung mittels Convolution – Rauschunterdrückung und Anomaly Detection mittels Autoencoder – Zeitreihenanalyse mit Long-Short-Term Memory Netzwerken (LSTM)
Lehrmittel	<ul style="list-style-type: none"> – Jupyter Notebooks (Python Skripts) – Ausgewählte Youtube-Videos im Selbststudium – Literaturempfehlung Nr. 15

8.11 Wahlfach Text Analytics und Natural Language Processing

Die maschinelle Gewinnung von Information aus Text, in unterschiedlichen Sprachen, in unterschiedlichen Formaten, häufig im Kontext von bestimmten Situationen welche nicht selbst im Text beschrieben sind, oder nur in vernetzten anderen Dokument vorkommen, ist äusserst komplex. Allerdings liegt auch ein riesiger Teil von Unternehmensinformation in Textform vor, beispielsweise Emails. Social Media Netzwerke sind ebenfalls riesige Informationsträger und werden immer intensiver ausgewertet.

Lernziele	Die Teilnehmenden lernen grundlegende Verfahren und Vorgehensweise zur Gewinnung von Information aus Texten kennen, beispielsweise für die Extraktion von spezifischen Informationen (z.B. genannter Personen oder Orte), die Klassifikation von Texten (z.B. Sentiment Analytics) und das Ranking von Dokumenten.
Themen und Inhalte	<ul style="list-style-type: none"> – Grundlegende Verfahren zur Text-Analyse (Natural Language Processing) – Klassifikation von Texten (z.B. Sentiment Analyse) – Praktische Anwendung obengenannter Verfahren mit Hilfe von Python und spezifischen Open Source-Bibliotheken auf ausgewählten Beispiel-Datensätzen
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript – Literaturempfehlung Nr. 16

8.12 Wahlfach Social Network Analytics

Soziale Netzwerke sind riesige Informationsträger, aber auch Instrumente für die Verbreitung und Diffusion von Information, für Prognosen und Forschungsfragen. Methodisch können soziale Netzwerke mit statistischen Verfahren, Graphentheorie, Text-, Bild- und Audio-Analysen beschrieben und analysiert werden. In der Informatik haben sich zahlreiche Datenbanksysteme und Programmibliotheken etabliert, welche eine einfache und effizienten Nutzung ermöglichen.

Lernziele	Die Teilnehmende können: <ul style="list-style-type: none"> – Die Begriffe der Sozial Netzwerkanalyse verstehen – Netzwerkdatenstrukturen manipulieren – Netzwerke visualisieren – Deskriptive Statistik auf globaler und lokaler Ebene rechnen – Statistische Modelle von dynamisch Netzwerke schätzen – Prozesse der Diffusion and Netzwerkentwicklung erklären – Erlernte Methoden in ihrem Arbeitsumfeld einsetzen
Themen und Inhalte	Methoden <ul style="list-style-type: none"> – Datenverwaltung und Visualisierung – Dynamische Modellierung und Diffusion Anwendungsbereiche <ul style="list-style-type: none"> – Epidemiologie – Wirtschaft/ Marketing – Sozialleben
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript – Literaturempfehlungen Nr. 17, 18, 19, 20

8.13 Wahlfach Recommender Systems

Lernziele	Sie kennen wichtige und erfolgskritische Faktoren beim Aufbau und Betrieb von Empfehlungssystemen und lernen anhand von konkreten Anwendungs- und Datenbeispielen die Funktionsweise von Empfehlungssystemen kennen.
Themen und Inhalte	<ul style="list-style-type: none"> – Was sind Empfehlungssysteme (Recommender Systems)? – Anwendungsfälle, Nutzen und Ziele – Ausgewählte Methoden und Algorithmen – Aufbau und Betrieb von Empfehlungssystemen – Projektbeispiele aus sozialen Medien, e-Commerce-Plattformen, Medienhäusern und Marktplätzen <p>Im Workshop wird anhand konkreter Daten (csv, R/Python) die Methodik, der Aufbau und die Herausforderungen an Empfehlungssysteme vermittelt.</p>
Lehrmittel	<ul style="list-style-type: none"> – Folien/Skript – Literaturempfehlung Nr. 21

8.14 Projektarbeit

Zielsetzung und Thema	<p>In der Projektarbeit (Semesterarbeit) bearbeiten die Teilnehmenden ein Projekt oder eine Fragestellung aus ihrer Firma, zum Thema Machine Learning.</p> <p>Die Semesterarbeit soll den ganzen Zyklus einer typischen ML-Aufgabe umfassen, von der Fragestellung bis hin zur Evaluation der Resultate. Dabei dürfen aber Schwerpunkte auf bestimmte Prozessschritte gelegt werden.</p> <p>Anstelle einer Projektarbeit in der Firma können auch Themen aus eigenem Interesse heraus definiert und bearbeitet werden.</p> <p>Auch Kaggle Competitions können bearbeitet werden, siehe https://www.kaggle.com/competitions in der Regel aus der Kategorie Research, Featured, Getting Started (mit Anpassungen).</p> <p>Das Arbeiten in Zweierteams ist ausdrücklich möglich und wünschbar.</p>
Ablauf	<p>Die Semesterarbeit umfasst ca. 90 Stunden Arbeit und beinhaltet folgende Meilensteine (siehe auch Zeitplan):</p> <ol style="list-style-type: none"> 1. In der Firma ein Thema suchen, und mit Vorteil einen Ansprechpartner / Betreuer in der Firma definieren. 2. Erstellen einer Projektskizze (Wordvorlage, 1 bis 2 Seiten) und einer Kurzpräsentation (Powerpoint, wenige Slides): <ol style="list-style-type: none"> a. Titel b. Umfeld c. Problemstellung d. Lösungsansatz (Vorgehen, Methoden) e. Name und Kontaktadressen der Gruppenmitglieder, und des Ansprechpartners / Betreuers in der Firma 3. Kurzpräsentation des Themas vor Dozentengremium, 5-10' Präsentation, 5-10' Diskussion, max. 15'. 4. Eventuell Überarbeitung der Projektskizze gemäss Feedback.

	<ol style="list-style-type: none"> 5. Zuordnung eines Experten durch die Schule. 6. Durchführung der Arbeit in eigener Terminplanung. 7. 2-3 Meetings mit dem Experten/Expertin (Durch Studierende organisiert) 8. Schlusspräsentation vor Klasse, Experte und Dozenten. 15' Präsentation, 15' Diskussion. 9. Abgabe des Berichtes an den Experten (per Email, auf Wunsch in Papierform) und den/die CAS-Verantwortliche.
Ergebnis und Bewertung	<p>Der Bericht ist in elektronischer Form als PDF-Dokument an den Betreuer zu schicken und auf der Moodle-Plattform zu hinterlegen.</p> <p>Bericht: ca. 20-30 Seiten, Source Code soweit notwendig für die Projektbeurteilung.</p> <p>Die Semesterarbeit wird nach folgenden Kriterien bewertet:</p> <ul style="list-style-type: none"> – Themeneingabe Projektskizze rechtzeitig und vollständig eingereicht. Themenpräsentation sorgfältig vorbereitet. Idee oder Aufgabe durchdacht und abgegrenzt, Quellen recherchiert, Rahmenbedingungen definiert, Teilziele priorisiert. – Methodik und Ausführung Gewählte Methode(n) systematisch und korrekt angewendet. Kreativ und agil in der Ausführung. Entscheidungen präzise begründet. – Ergebnis Nachvollziehbares und dokumentiertes Ergebnis. Aufgabenstellung erfüllt. Ergebnisse validiert, getestet, verifiziert. Vergleich von Zielsetzung und Ergebnis vorgenommen. Learnings und Ausblick vorhanden. – Bericht und Dokumentation Vollständig und verständlich. Rechtschreibung korrekt. Kapiteleinteilung sinnvoll. Angemessene Darstellung. Grafiken auf das Wesentliche reduziert und beschriftet. – Schlusspräsentation Roter Faden, logisches Vorgehen, klare Aussagen. Identifikation mit dem Thema spür- und erkennbar. Professionelle Präsentationstechnik, Zeitvorgaben genutzt und eingehalten. Fragen präzise und sicher beantwortet.
Vertraulichkeit	<p>Semesterprojekte werden sinngemäss wie Master Thesen behandelt, d.h. grundsätzlich als nicht-öffentliche Projekte. Es steht ein kostenloses Standard NDA der Schule zur Verfügung. Individuelle Vereinbarungen sind kostenpflichtig.</p>

9 Kompetenznachweis

Für die Anrechnung der 12 ECTS-Punkte ist das erfolgreiche Bestehen der Qualifikationsnachweise (Prüfungen, Projektarbeiten) erforderlich, gemäss folgender Aufstellung:

Kompetenznachweis	Gewicht	Art der Qualifikation	Erfolgsquote Studierende
Gesamtprüfung Fragen zu allen Themengebieten	4	Schriftliche Einzelprüfung, Open Book, Fragen zu allen Themengebieten	0 - 100 %
Projektarbeit	6	Semesterarbeit	0 - 100 %
Gesamtgewicht / Erfolgsquote	10		0 - 100 %
ECTS-Note			A - F

Jeder Studierende kann in einem Kompetenznachweis eine Erfolgsquote von 0 bis 100% erreichen. Die gewichtete Summe aus den Erfolgsquoten pro Thema und dem Gewicht des Themas ergibt eine Gesamterfolgsquote zwischen 0 und 100%. Die Gesamterfolgsquote wird in eine ECTS Note A bis E umgerechnet, gemäss Studienreglement. Weniger als 50% Gesamterfolgsquote ergibt eine ungenügende Note F.

10 Lehrmittel

Die nachfolgend aufgeführten Lehrmittel sind wesentlich für das Lernen während des geführten Unterrichtes. Sie sind durch die Studierenden zu beschaffen.

Nr	Titel	Autoren	Verlag	Jahr
1.	Data Mining	Ian H. Witten Eibe Frank Mark A. Hall Christopher J. Pal	Morgan Kaufmann	2016
2.	Introduction to Machine Learning with Python	A. Müller S. Guido	O'Reilly Media	2016
3.	Machine Learning in Action	P. Harrington	Manning Publications	2012
4.	R in Action: Data Analysis and Graphics with R	Robert I. Kabacoff	Manning	2011
5.	R Cookbook, Proven Recipes for Data Analysis, Statistics, and Graphics, 2nd ed.	Paul Tretor	O'Reilly	2019

Nr	Titel	Autoren	Verlag	Jahr
6.	Applied Predictive Modeling	Kjell Johnson Max Kuhn	Springer	2018
7.	Feature Engineering for Machine Learning	Alice Zheng Amanda Casari	O'Reilly	2018
8.	allfällige weitere werden bei Kursbeginn bekannt gegeben			
9.	Offizielle Python Webseite		Englisch https://docs.python.org/3/tutorial/index.html Deutsch https://py-tutorial-de.readthedocs.io/de/python-3.3/	
10.	Learning to Program with Python	Richard L. Halterman	https://www.cs.uky.edu/~keen/115/Haltermanpythonbook.pdf	2011
11.	Python for everybody	Charles R. Severance	http://do1.dr-chuck.com/pythonlearn/EN_us/pythonlearn.pdf	2009
12.	Machine Learning in Action, Chapter 10 + 11	Peter Harrington	Manning Publications	2012
13.	Large-Scale Collaborative Filtering for the Netflix Prize	Zhou et al.	Proceedings of AAIM. 2008	2008
14.	Data Smart: Using Data Science to Transform Information into Insight, Chapter 2	John W. Foreman	John Wiley & Sons	2013
15.	Deep Learning Book	Ian Goodfellow Yoshua Bengio Aaron Courville	MIT Press	2016
16.	Natural Language Processing with Python	Steven Bird Ewan Klein Edward Loper	O'Reilly Media	2009
17.	Mining the Social Web	Matthew Russell	O'Reilly Media	2013
18.	Statistical Analysis of Network Data with R.	Kolaczyk and Csardi		2015
19.	Social Network Analysis	Wasserman and Faust		1994
20.	Analyzing Social Networks	Borgatti, Everett and Johnson		2013
21.	Recommender Systems	© Dietmar Jannach Markus Zanker Alexander Felfernig Gerhard Friedrich	Cambridge University Press.	2011

11 Dozierende

Vorname Name	Firma	E-Mail
Jürgen Vogel	Berner Fachhochschule	juergen.vogel@bfh.ch
Werner Dähler	Staatssekretariat für Wirtschaft	werner.daehler@bfh.ch
Marcus Hudritsch	Berner Fachhochschule	marcus.hudritsch@bfh.ch
Matthias Dehmer	Universität der Bundeswehr München	matthias.dehmer@unibw.de
Niklaus Johner	Uni Basel, 4teamwork	niklaus.johner@gmail.com
Debra Hevenstone	Berner Fachhochschule	debra.hevenstone@bfh.ch
Romeo Kienzler	IBM	romeo.kienzler@ch.ibm.com
Salvatore Dino	Microsoft	salvatore.Dino@microsoft.com

- + Weitere Dozenten und Referenten aus Wirtschaft und Lehre
- + Weitere Experten und Betreuer für die Projektarbeit

12 Organisation

CAS-Leitung:

Arno Schmidhauser

Tel: +41 31 84 83 275

E-Mail: arno.schmidhauser@bfh.ch

CAS-Administration:

Andrea Moser

Tel: +41 31 84 83 211

E-Mail: andrea.moser@bfh.ch

Dokumenteninformation

Study Guide CAS Practical Machine Learning
Stand: 3. September 2019

Dieser Study Guide gilt für die Publikation ab Herbstsemester 2019.

Während der Durchführung des CAS können sich Anpassungen bezüglich Inhalten, Lernzielen, Dozierenden und Kompetenznachweisen ergeben. Es liegt in der Kompetenz der Dozierenden und der Studienleitung, aufgrund der aktuellen Entwicklungen in einem Fachgebiet, der konkreten Vorkenntnisse und Interessenslage der Teilnehmenden, sowie aus didaktischen und organisatorischen Gründen Anpassungen im Ablauf eines CAS vorzunehmen.

Berner Fachhochschule
Technik und Informatik
Weiterbildung
Wankdorffeldstrasse 102
CH-3014 Bern

Telefon +41 31 848 31 11
Email: office.ti-be@bfh.ch

bfh.ch/ti/weiterbildung
ti.bfh.ch/cas-pml