# **CAS Applied Data Science Study Plan**

-ð

9



b UNIVERSITÄT BERN

15.09.2022	The Applied Data Science Course is a university study program leading to a "Certificate of Advanced Studies in Applied Data Science" awarded by the University of Bern as laid out in the regulations of the Faculty of Science of the University of Bern for the Studiengänge in Extended Intelligence of 9. Dezember 2021.
	1. Course objectives
About	In most disciplines, the amount of available data is growing rapidly. Han- dling and processing the data remains a challenge and the demand for data science expertise is increasing dramatically across all domains. This CAS offers internal and external students, researchers, and employees the oppor- tunity to complement their domain-specific core competences with a formal deepening and broadening of their applied data science knowledge and skills. The format is designed to align with the participants' main study and professional activities. The teaching and learning approaches are team and discussion oriented and designed to develop practical competency.
Competence objectives	The course competences are developed in six modules. At the end graduates will be:
	<i>1.</i> familiar with different data sources and data types, and able to develop data management plans;
	2. able to describe, extract and present scientific knowledge from data us- ing statistical methods;
	3. able to process data with machine learning tools and methods;
	4. familiar with best practices for data management, analytics, and science;
	5. able to analyse and communicate data science challenges and use a wide range of data science tools and methods;
	$\boldsymbol{\delta}$ able to train and apply deep neural networks for a range of tasks;
	7. able to perform data science projects with machine learning.
	2. Duration, outcomes and objectives of the modules
Duration and scope	The CAS consists of 18 days of courses (126 hours attendance) and 16 ECTS credit points (total effort about 480 hours) whereof 4 ECTS result from the CAS project work. Attendance at one introductory <i>About the CAS Applied Data Science</i> meet- ing before course admission is mandatory.

# Data acquisition and management

In which participants learn to understand data sources and types, how to design data flows and models, and manage and visualise data algorithmically

ECTS	2 ECTS credit points (incl.	Duration	3  days = 21		
	self-study and report)		hrs.		
			attendance		
Assess-	Conceptual Design Report	Req. at-	80%		
ment		tendance			
Learning	Graduates will				
outcomes	• know data sources, types,	sizes, storage	e systems, com-		
	mon infrastructures and ac	equisition tec	hniques;		
	• be able to analyse data flo	ws;			
	• be able to develop concep	tual, logical a	and physical		
	data models;		_		
	know how to use database	s and MySQ	L;		
	• be able to design data mar	nagement pla	ns;		
	• will be able to collect, manage and visualise data.				
Learning	Definitions for data and science				
objectives	• Types of sensors and data				
	Quality of data				
	Infrastructures for data				
	Data flows and models				
	Algorithmic data management, filtering, cleaning				
	and visualization				
-	• MySQL				
Learning	Online platform				
and teach-	• Lectures				
ing meth-	Hands-on tutorials				
ods	Discussions				
	Report writing				
Prerequi-	None				
sites					
Teaching	English				
language	-				

## Module 2 Statistical Inference for Data Science

In which we recall or become familiar with typical statistical concepts for describing and analysing data.

ECTS	2 ECTS credit points (incl. self-study and project)	Duration	3 days = 21 hours	
			attendance	
Assess-	Poster presentation	Req. at-	80%	
ment		tendance		
Learning	Graduates will/can			
outcomes	<ul> <li>know the importance of statistical inference for data science and where to apply it;</li> <li>understand the basic theoretical concepts;</li> </ul>			

	<ul> <li>apply the theoretical concepts with software on data;</li> <li>draw scientific conclusions from statistical analysis results.</li> </ul>
Learning objectives	<ul> <li>Probabilities and distributions</li> <li>Significance and p-values</li> <li>Estimation paradigms (least squares, maximum likelihood)</li> <li>Uncertainties (standard errors, confidence intervals)</li> <li>Basic regression</li> <li>Hypothesis testing</li> </ul>
Learning and teach- ing meth- ods	<ul> <li>Online platform</li> <li>Lectures</li> <li>Inverted classroom with hands-on tutorials</li> <li>In-person classes for discussions, feedback and deepening of knowledge</li> <li>Presentation of results in poster session</li> </ul>
Prerequi- sites	<ul> <li>Basic experience with a programming language providing statistical libraries (Python, R,)</li> <li>Knowledge in mathematics and programming at the level of a university introduction lecture</li> </ul>
Teaching language	English

#### Module 3

 $\geq^{a}$ 

ţ

# Data Analysis with Machine Learning

In which participants le	earn about	standard	machine	learning to	echniques
and how to apply them.					

ECTS	2 ECTS credit points (incl.	Duration	3  days = 21  hrs		
	self-studies and project)		attendance		
Assess-	Presentation	Req. at-	80%		
ment		tendance			
Learning	Graduates will				
outcomes	• know the basic concepts o	f machine le	arning: training,		
	testing, over/ and underfit	ting, perform	ance measures;		
	• can apply software to perf	òrm linear re	gression, deci-		
	sion trees and random for	est, neural ne	tworks (ML		
	software is accessed via P	ython.			
Learning	• Machine learning overview, classification and regres-				
objectives	sion				
	Linear model and logistic regression				
	• Trees and forests				
	PCA and embeddings				
	Neural networks				
Learning	Online platform				
and teach-	• Lectures				
ing meth-	Hands-on tutorials				
ods	Project work				
	Presentations				
Prerequi-	Python programming experience				
sites					

	• Knowledge in mathematics and programming at the level of an introductory lecture
Teaching language	English

Module 4

# AI philosophy, ethics and peer exchange

In which participa	nts reflect upon	philosophical	and ethical	aspects of Arti-
ficial Intelligence	(AI) and share (	and discuss da	ta science ex	xperiences

ECTS	2 ECTS credit points (incl.	Duration	3 days = 21	
	self-studies		hours attend-	
			ance	
Assess-	presentation	Req. at-	80%	
ment		tendance		
Learning	Graduates will/can			
outcomes	• know the philosophical co	nceptions of	AI;	
	• know main moral challeng	ges related to	AI and discuss	
	solutions from the perspec	tive of ethics	3;	
	• use distributed VCS softw	are and platf	orms;	
	• present and discuss data set	cience topics		
Learning	• History of AI and philosophical conceptions of AI			
objectives	• Ethical challenges due to AI in the perspective of			
	• Ethical chantenges due to	At in the pers	spective of	
	ethical theories, machine e	ethics		
Learning	Online platform			
and teach-	Seminar discussions			
ing meth-	• Oral presentations			
ods	•			
Prerequi-	Data analysis experience			
sites				
Teaching	English			
language				

### Module 5

# **Best Practices for Data Science**

In which best practices and tools for data and code management, resource usage and collaboration are presented, discussed and applied for the CAS projects

ECTS	2 ECTS credit points (incl.	Duration	3  days = 21		
	self-study + module work)		hours atten-		
			dance		
Assess-	Project presentation	Req. at-	80%		
ment		tendance			
Learning	Graduates will				
outcomes	know best practices for scientific computing;				
	• be able to use distributed Version Control Software				
	(VCS);				
	• know basic cyber security challenges;				
	• know legal aspects regarding intellectual property and				
	licensing;				

	• be able to document and publish software projects with VCS and websites
	Veb and websites.
Learning	• Best practices for scientific computing
objectives	Collaborative distributed version control, code review
	(course tool is Git and related platforms)
	Basic cyber security
	Intellectual property and licenses
	Documentation
Learning	Online platform
and	• Lectures
Teaching	Hands-on tutorials
methods	Inverted classroom
	Project work
	• Presentations
Prerequi-	Python programming experience
sites	• Knowledge in mathematics and programming at the
	level of an introductory lecture
Teaching	English
language	

Module 6

**Deep Learning** In which participants learn about deep learning techniques and how to ap-ply them

piy inem.				
ECTS	2 ECTS credit points (incl.	Duration	3  days = 21	
	self-studies and project)		hours atten-	
			dance	
Assess-	Project presentation	Req. at-	80%	
ment		tendance		
Learning	Graduates will			
outcomes	• understand different mode	ls for superv	ised, unsuper-	
	vised and reinforced learn	ing;		
	• be able to apply software	for deep learn	ning (course	
	software is TensorFlow);			
	• be able to train, tune and a	issess deep n	etworks.	
Learning	• TensorFlow			
objectives	• Training of deep models			
-	• Tuning of deep models			
	Model performance assess	sment		
Teaching	Online platform with court	se material		
methods	• Theoretical and applied le	ctures		
	Hands-on tutorials	••••		
	Project work with machine learning software			
	Project report			
	Oral presentations			
Proroqui	Oral presentations     Duthen programming and data handling			
sites	Python programming and data nandling			
51105	Basic linear algebra, analysis, statistics and machine			
Teaching	English			
language				

### Project Work CAS Project Work

In which participant teams perform and present a comprehensive data science project based on all CAS modules.

ECTS	4 ECTS credit points	Duration	120 hours
Assess-	Written report	Req. at-	No attendance
ment		tendance	
Learning	Graduates will		
outcomes	• be able to perform and present a data science project with machine learning.		
Learning objectives	• Selection of objectives from all modules		
Teaching	Project work with remote supervision		
methods	Written report		
Prerequi-	• Module 1 - 6		
sites			
Teaching language	English		

#### 3. Performance assessment

Assessment

The performance assessment includes 80% attendance and passed performance assessments from all modules and the CAS project report.

- a) For Module 1 a written data science concept design report is assessed.
- b) For Module 2 a poster presentation is assessed.
- c) For Module 3, 4, 5 and 6 oral presentations are assessed.

For the CAS project a written report is assessed.

#### 4. Final regulations

Entry into force The present plan shall enter into force on 01.11.2022.

15.09.2022 Released by the program management:

mistione

wella

Prof. Dr. Christiane Tretter (Chair)

06.10.2022 Adopted by the Faculty of Science, University of Bern

The Dean

Prof. Dr. Mardo Herwegh