



## Data Science: Applied Text Mining (S42) 25 – 29 July 2022

Course Director: Dr. Ayoub Bagheri  
E-mail: ms.summerschool@uu.nl

Saturday 23 & Sunday 24 July 2022		
Time	Activity	Description
12.00 – 18.00	Key pick up	<i>You will find the exact key pick up location in the pre-departure information, which becomes available after you have paid the course fee.</i>

Day	Time	Type	Description	Location
Monday	09:00 – 10:30	Lecture	Introduction to text mining Preprocessing text Feature extraction	UTRECHT SCIENCE PARK, EXACT LOCATION TO BE ANNOUNCED
	10:45 – 11:45	Computer Lab	Python in Google Colaboratory	
	11:45 – 12:30	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	
	14:00 – 15:15	Lecture	Text classification: - Binary classification - Multi-class classification	
	15:30 – 16:30	Computer Lab	Document-term matrix Text classification & its evaluation	
	16:30 – 17:00	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	



Day	Time	Type	Description	Location
Tuesday	09:00 – 10:30	Lecture	Feature selection in text	UTRECHT SCIENCE PARK, EXACT LOCATION TO BE ANNOUNCED
	10:45 – 11:45	Computer Lab	Filter-based feature selection Embedded feature selection Dimension reduction	
	11:45 – 12:30	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	
	14:00 – 15:30	Lecture	Text clustering & topic modeling	
	15:45 – 16:30	Computer Lab	K-Means clustering Topic modeling	
	16:30 – 17:00	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	

Day	Time	Type	Description	Location
Wednesday	09:00 – 10:30	Lecture	Word embedding Distributional hypothesis	UTRECHT SCIENCE PARK, EXACT LOCATION TO BE ANNOUNCED
	10:45 – 11:45	Computer Lab	Word embeddings Pre-trained embeddings FastText	
	11:45 – 12:30	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	
	14:00 – 15:30	Lecture	Feed-forward neural networks Recurrent neural networks	
	15:45 – 16:30	Computer Lab	Neural networks The embedding layer Pretrained word embeddings	
	16:30 – 17:00	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	



Day	Time	Type	Description	Location
Thursday	09:00 – 10:30	Lecture	Convolutional Neural Networks Transformers	UTRECHT SCIENCE PARK, EXACT LOCATION TO BE ANNOUNCED
	10:45 – 11:45	Computer Lab	Recurrent neural networks Convolutional Neural Networks RNN vs CNN Hyperparameter Optimization	
	11:45 – 12:30	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	
	14:00 – 15:30	Lecture	Text mining applications	
	15:45 – 16:30	Computer Lab	Applications and pipeline	
	16:30 – 17:00	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	

Day	Time	Type	Description	Location
Friday	09:00 – 10:30	Lecture	Sentiment analysis	UTRECHT SCIENCE PARK, EXACT LOCATION TO BE ANNOUNCED
	10:45 – 11:45	Computer Lab	Sentiment classification and analysis	
	11:45 – 12:30	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	
	14:00 – 15:30	Lecture	Responsible text mining Bias and fairness in text mining	
	15:45 – 16:30	Computer Lab	Neural Machine Translation	
	16:30 – 17:00	Plenary Discussion	Students and teachers discuss and present their solutions to the computer lab	

For information about the Social Programme, please visit the [Utrecht Summer School website!](#)