

LECTURE / SESSION¹ PLAN

Name Lecturer: Sofia Gil-Clavel	Date session: 9	Expected number of students: 20
Course title	R-Workshop	
Topic of lecture or session	Machine Learning for Text	
Situational factors (e.g. group size, prior knowledge, expected motivation)	The class lasts 3hrs. The students are in different career stages, from PhD students to professors. The students have different backgrounds, from qualitative researchers that have never used R to quantitative researchers that want to learn new topics or move out of SPSS/STATA. The students use different Operating Systems (Windows, Mac, or Linux).	
Intended learning outcomes of this session	At the end of the session the students will be able to: <ol style="list-style-type: none">1. Situate statistical learning in the machine learning field.2. Explain the differences between supervised and unsupervised learning.3. Transform and prepare text for machine learning.4. Perform and visualize topic modelling.5. Situate text classification in the field of statistical learning.6. Set-up a pipeline for text classification.	
Learning material (book, chapters, ...)	The lecture is based on the books: <ul style="list-style-type: none">• Robinson, Julia Silge and David. <i>Welcome to Text Mining with R / Text Mining with R</i>. Accessed March 12, 2025. https://www.tidytextmining.com/.	

¹ A session means a teaching and learning session for university students in the bachelor or master. It can be a lecture, a seminar or a specific type of educational meeting with a group of students that is relevant for the discipline where the lecturer can demonstrate a whole range of teaching skills. It means that in one session each of the six didactic elements appears at least once. Your lesson plan should show that you are using a powerful learning environment and that students are activated.

	<ul style="list-style-type: none"> Silge, Emil Hvitfeldt and Julia. <i>Supervised Machine Learning for Text Analysis in R</i>. Accessed May 5, 2025. https://smltar.com/.
Media, equipment, tools	<p>The students use their own laptops.</p> <p>The teacher needs access to a projector and a whiteboard.</p>
Preparation for students	<p>The students have access to the slides and codes before the class:</p> <p>https://github.com/SofiaG11/R_Course/tree/master/R4SocialScientists/Session9_MachineLearning4Text</p>

Time ² (min.)	Didactic element (goal) ³ and topic ⁴	What the teacher does ⁵ (teacher activity)	What students do (learner activity)	Evaluation ⁶ (feedback/assessment)
10	<ul style="list-style-type: none"> Explain what machine learning is. 	<ul style="list-style-type: none"> The teacher uses the slides to introduce the topic: Machine Learning. The teacher opens RStudio and opens the already written script that will be used during the class. 	<ul style="list-style-type: none"> Before telling each element the students learned, the teacher waits some seconds for the students to fill out the information out loud. The students open and follow the steps that the teacher is explaining. 	<ul style="list-style-type: none"> The students will use these concepts during the workshop. So, the teacher will detect when a student confuses them. Based on these confusions the teacher will be able to correct the student.
10	<ul style="list-style-type: none"> Explain the difference between 	<ul style="list-style-type: none"> The teacher uses the slides to explain the differences between 	<ul style="list-style-type: none"> The students passively digest what the teacher is explaining. 	<ul style="list-style-type: none"> The students will use these concepts during the workshop. So, the teacher will detect

² Indicate the planned duration in minutes

³ State the number(s) of the relevant ILO at each didactic element

⁴ Only key words

⁵ Specify the type of activity and write down also the questions that you prepared to ask in the session

⁶ Specify the type of evaluation, i.e. the way in which you assess if the objective(s) has/have been achieved

	supervised and unsupervised learning.	supervised and unsupervised machine learning. <ul style="list-style-type: none"> The teacher gives a quick overview of the topics that will be learned today. 		when a student confuses them. Based on these confusions the teacher will be able to correct the student.
30	<ul style="list-style-type: none"> Transform and prepare text for machine learning. 	<ul style="list-style-type: none"> The teacher uses the slides to motivate and explain why it is important to transform text into numbers. The teacher uses the slides to explain how to do it. 	<ul style="list-style-type: none"> The students passively digest what the teacher is explaining. The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides. 	<ul style="list-style-type: none"> The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.
10	<ul style="list-style-type: none"> Break 	<ul style="list-style-type: none"> 		<ul style="list-style-type: none">
40	<ul style="list-style-type: none"> Perform and visualize topic modelling. 	<ul style="list-style-type: none"> The teacher uses the slides to motivate and explain topic modelling. The teacher uses the slides to explain how to perform and interpret Latent Dirichlet Allocation. 	<ul style="list-style-type: none"> The students passively digest what the teacher is explaining. The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides. 	<ul style="list-style-type: none"> The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.

20	<ul style="list-style-type: none"> Situate text classification in the field of statistical learning 	<ul style="list-style-type: none"> The teacher uses the slides to explain the differences between supervised and unsupervised machine learning for text. The teacher gives a quick overview of the topics that will follow during the lecture. 	<ul style="list-style-type: none"> The students passively digest what the teacher is explaining. 	<ul style="list-style-type: none"> The students will use these concepts during the workshop. So, the teacher will detect when a student confuses them. Based on these confusions the teacher will be able to correct the student.
50	<ul style="list-style-type: none"> Set-up a pipeline for text classification 	<ul style="list-style-type: none"> The teacher uses the slides to motivate and explain how to set-up the pipeline. 	<ul style="list-style-type: none"> The students passively digest what the teacher is explaining. The students independently start filling out the missing parts of the R-script. This is based on what the teacher explained using the slides. 	<ul style="list-style-type: none"> The teacher walks around the classroom to check on the students and provide feedback when something is not working on their computers. When the teacher detects a common error, then the teacher uses the whiteboard to clarify.