



# NER 2015 BCI Challenge

NICOLAS BERBERICH

ANDREAS WIEDERMANN

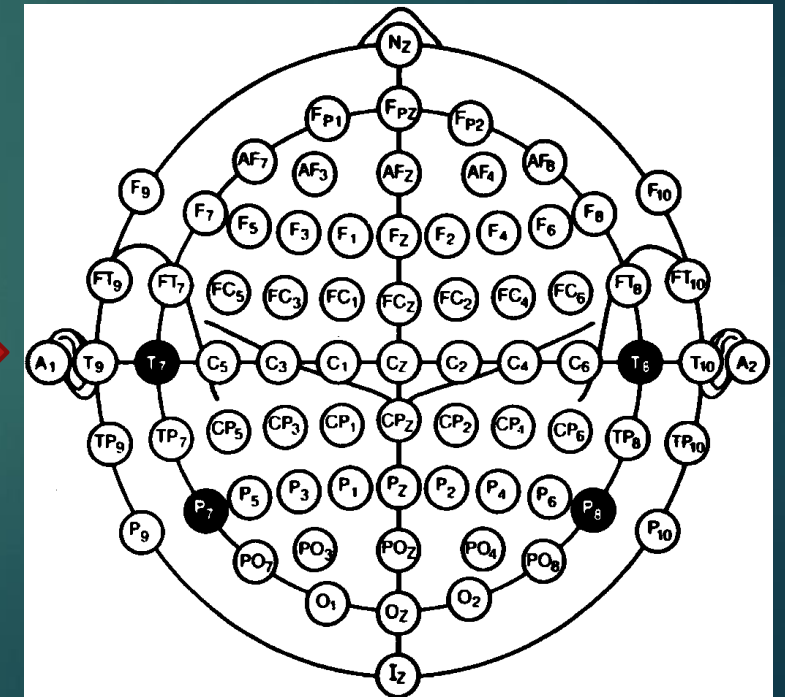
CLAAS BRÜß

# Outline

- ▶ The Challenge
- ▶ Course Relevance
- ▶ Data Format and Preprocessing
- ▶ Machine Learning Schemes
- ▶ Result Comparision
- ▶ Ranking and Outlook

# The Challenge

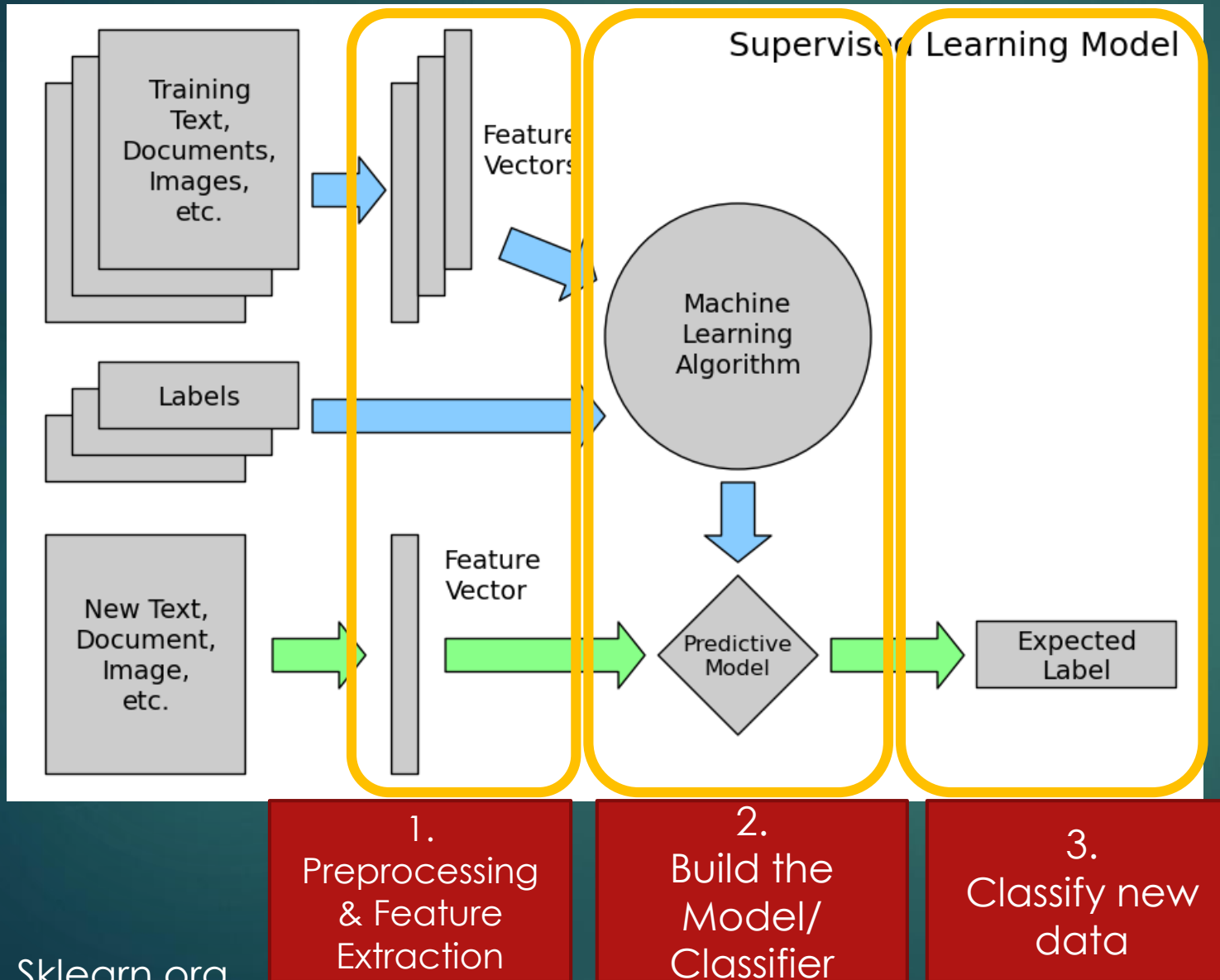
- ▶ This particular Kaggle challenge has already been archived and was posed as part of the IEEE Neural Engineering Conference 2015 (NER2015) with the goal to flag errors in brain-computer interface responses through analysis of EEG data in the test subjects.



# Course Relevance

- ▶ EEG Data mapping brain activity
- ▶ Due to data complexity machine learning algorithms are applied some of which are brain inspired such as certain neural networks
- ▶ The experiment is based on a BCI setup and is geared towards improvements in BCI setups in order to make a more general application of BCI possible

# Classification Process



# Classification process

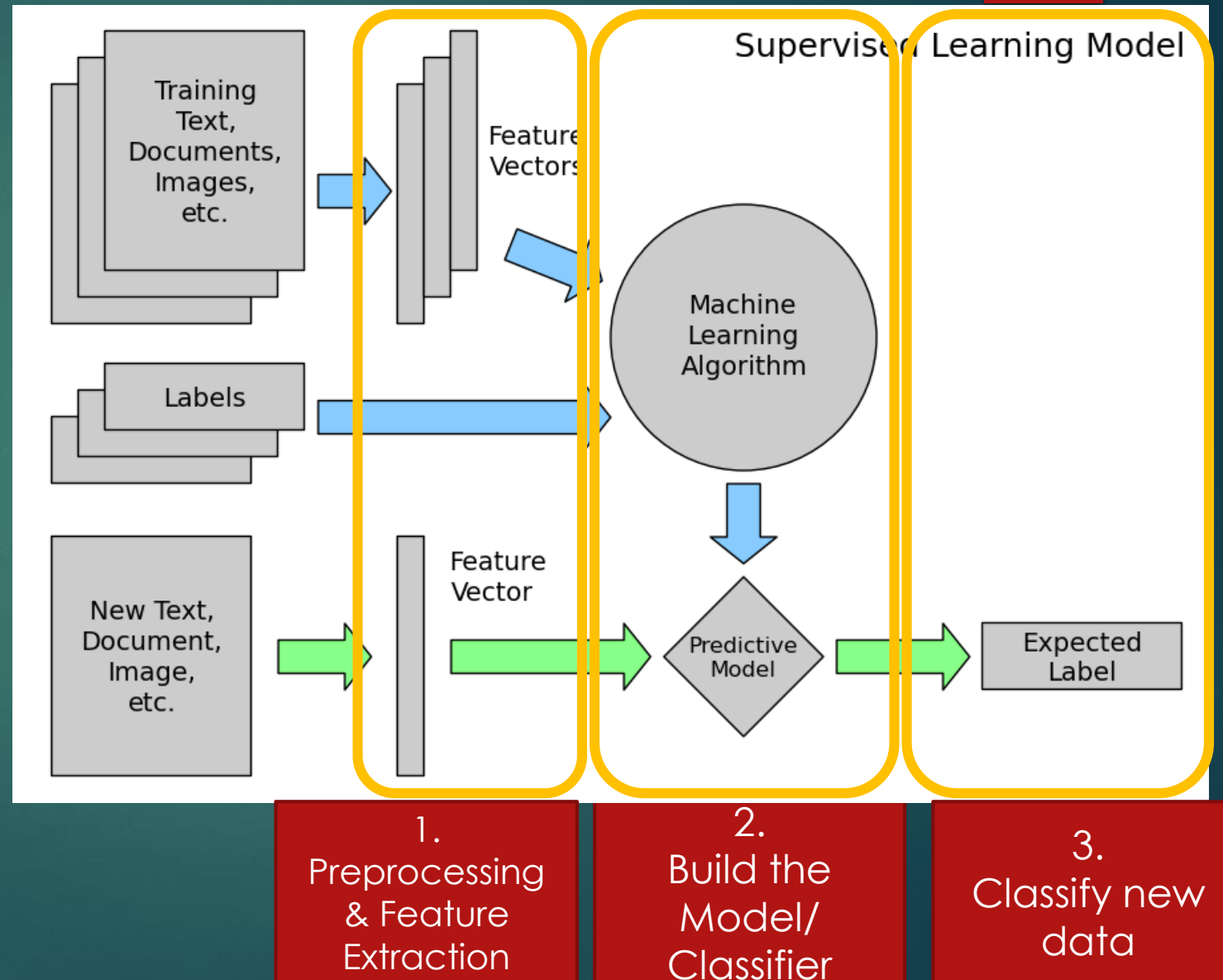
Data Format:

EEG data:

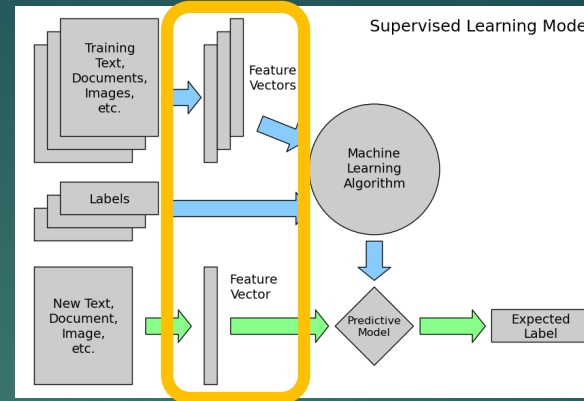
- 16 test subjects, 5 sessions per person = 80 csv files
- 56 electrodes was measured at 200 Hz
- The timing of feedback events was labeled

→ (= 8.7 GB)

	A	B	C	D	E
1	Time,Fp1,Fp2,AF7,AF3,AF4,AF8,F7,F5,F3,F1,Fz,F2,F4,F6,F8,FT7,FC5,FC3,FC1,FCz,				
2	0.000000,-449.704315,-99.462274,436.308045,-71.140183,115.558084,284.0				
3	0.005000,-731.824275,-164.517846,693.700552,-111.532378,177.992746,445				
4	0.010000,-642.515558,-132.375471,628.258165,-90.905004,172.435525,410				
5	0.015000,-701.644987,-161.379741,651.696937,-112.723206,162.286167,417				
6	0.020000,-656.383508,-143.265979,651.519421,-91.885759,177.981647,419				
7	0.025000,-686.968198,-154.415145,649.957807,-107.205403,161.112334,411				
8	0.030000,-661.975139,-142.455586,658.623306,-97.123994,168.942177,424				
9	0.035000,-686.647566,-158.936459,642.423443,-118.191821,155.823151,412				



# 1. Preprocessing



Preprocessing:

- ▶ Feedback related 1.3 second long chunks of EEG data cut out of complete data set
- ▶ Application of Butterworth filter with a passband of 1-40 Hz (alpha activity 8-13 Hz, beta activity 13-31 Hz)
- ▶ Amplitude Average over 1.3 seconds of the Butterworth filtered data



# 2. Machine Learning Algorithms

## sklearn based implementations

### Random Forests:

- ▶ Number of trees  $n = 500$  or  $1000$
- ▶ best result 0.697, Channel 39,  $n = 500$

### Gradient Boosting Machine:

- ▶ Number of estimators  $n = 500$  or  $1000$
- ▶ a single regression tree fit on the negative gradient of binomial deviance loss function
- ▶ Best result 0.644, Channel 39,  $n = 500$

### Support Vector Machines:

- ▶ Precision requirement  $C = 5$
- ▶ runtime issues with training ( $>15h$ )
- ▶ best result 0.636, Channel 39, downsampled 5x,  $n = 500$

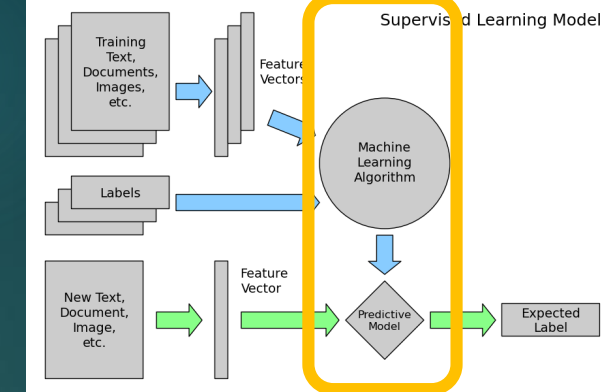
## EEGlearn based implementations

### Fully Connected Neural Network:

- ▶ Linear sequential stack, epochs = 3000
- ▶ Best result with cross referenced train data 0.84
- ▶ Corresponding test result 0.5 (random)

### Convolutional Neural Network:

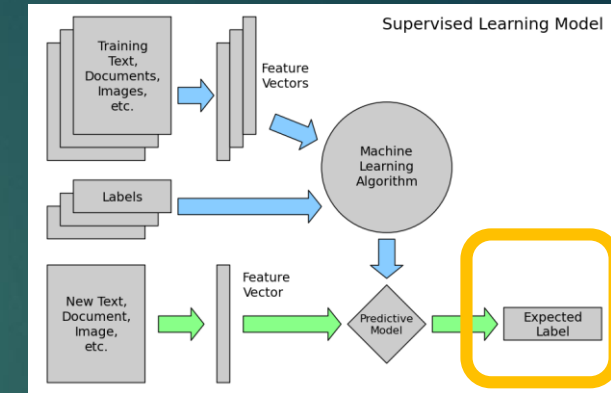
- ▶ Require images as input
- ▶ Alpha, Beta, Theta bands are being averaged for this
- ▶ Result 0.542, Epochs = 3000







# 3. Classification Results Comparision

Feature	Classifier	Channel, Electrode	Parameter	Result acc
Raw Time Series	Gradient Boosting	channel 46	n=500	0.573
		channel 46	n=1000	0.58
		channel 40	n=500	0.582
	Random Forest	channel 39	n=500	0.609
		channel 46	n=500	0.577
		channel 46	n=1000	0.574
		channel 40	n=500	0.595
		channel 39	n=500	<b>0.647</b>
		channel 40	n=1000	0.639
Filtered Time Series	Gradient Boosting	channel 46	n=500	0.584
		channel 39	n=500	0.644
	Random Forest	channel 46	n=500	0.653
		channel 46	n=1000	0.678
		channel 39	n=500	<b>0.697</b>
		channel 39	n=1000	0.684
		channel 39	n=1500	0.682
	Support Vector Machine	channel 39	5xdown, n200	0.623
		channel 39	5xdown, n500	0.636
		channel 39	10xdown, n200	0.630
		channel 39	10xdown, alle	0.624
	Fully Connected Neural Network	channel 39	Epochs = 3000	0.5 ( <b>0.84</b> cross reference)
Filtered Average	Gradient Boosting	channel 39	n=500	0.518
	Random Forest	channel 39	n=500	<b>0.533</b>
Averaged three channel image	Convolutional Neural Network	Averaged theta, alpha, beta	Epochs = 3000	<b>0.542</b>



- ▶ Random Forest Approach delivers best results in all categories
- ▶ Prior knowledge about EEG data motivated Butterworth Filter application
- ▶ Filtering showed positive impact
- ▶ Averaging showed negative impact

# Ranking and Outlook

12	↑33	ThibaultV	<a href="#">0.70152</a>	20	<a href="#">Mon, 26 Jan 2015 10:57:45 (-5.5d)</a>
13	↑136	A.M.	<a href="#">0.69865</a>	11	<a href="#">Wed, 11 Feb 2015 23:39:57 (-13.5d)</a>
-		<b>Erzlearner</b>	<b>0.69659</b>	-	<b>Tue, 05 Jul 2016 22:36:18</b> <b>Post-Deadline</b>
<b>Post-Deadline Entry</b> If you would have submitted this entry during the competition, you would have been around here on the leaderboard.					
14	↑29	os	<a href="#">0.69468</a>	12	<a href="#">Sat, 21 Feb 2015 18:02:02 (-3.9d)</a>
15	↑12	 Black Swan (NER) 	<a href="#">0.69364</a>	39	<a href="#">Tue, 20 Jan 2015 01:02:22 (-6.9d)</a>

- ▶ Challenge felt impactful
- ▶ Motivation to work on future challenges
- ▶ Future projects in this subject area