

# Updated Project Proposal

## Project Timeline

When	Who	What	Time Spent
Week 39	Akash and Lisa	Find datasets, do research on different kinds of dataset	20h
	Chris and Lisa	Extract and label relevant data from online datasets	10h
	Niko and Chris	Find classifier libraries	10h
Week 40	Chris and Lisa	Extract and label relevant data from online datasets.	10h
	Akash and Niko	Create code skeleton (interfaces)	13h
	Akash and Chris	Construct a dictionary	16h
	Niko and Lisa	Implement bag of word	16h
	Akash and Lisa	Integrate all components	16h
<b>Friday 2<sup>nd</sup> October Milestone: Implement general text analyser prototype</b>			
Week 41	Niko and Chris	Define Tests	6h
	Akash and Lisa	Customize prototype for project specific data	18h
	Akash and Lisa	Test different datasets	16h
	Niko and Chris	Compare classifiers	16h
	Akash and Niko	Benchmarking	8h
<b>Friday 9<sup>th</sup> October Milestone: Implement data specific prototype.</b>			
Week 42	Chris and Lisa	Compare feature extraction	16h
	Akash and Niko and Lisa and Chris	Refactoring and optimization	16h
	Niko and Akash	Project report	20h
		Backup Time	8h
<b>Thursday 15<sup>th</sup> October Milestone: Project code &amp; report delivery.</b>			
Week 43	Akash, Nico, Lisa and Chris	Prepare presentation	16h
<b>Tuesday 20<sup>th</sup> October Milestone: Project Presentation</b>			

## Comments on Updates

There have been some significant changes in the project timeline and estimation of resource allocation. These changes are mainly due to underestimated time required for certain tasks such as finding, retrieving and labelling datasets in appropriate formats. Other tasks such as finding tools for feature extraction was overestimated and has thus been adjusted.

To this date we have a working prototype which works for general datasets and should now be customized to work for our specific training data. As reflected in the timeline some tasks have been prioritized based on the experience gained during the first week of research. We have done this

according to the setup explained in the first draft of the project proposal. It turned out that this was already accurate and we did not had to make changes to it.