**Introduction** (quarter of a page)

* Research question:
* Which variables contribute to subscription of term deposits?
* Relevance:
* Helps banks in deciding their marketing strategy, and which groups to reach out to
* Method used:

1. Support vector machine
2. Cross validation (with adjusted rand index)
3. Linear and non-homogenous kernels (to model non-linear effects)

* Results:

1. Which variables contribute the most?
2. Which model, in general, is best at predicting?

**Data** (half a page) (Joyce)

* Describe the following:

1. Where does the data come from?

* Time, place (Portugal), institution (bank) (Cite that it is from the UCI machine learning repository)
* Justify why we work with sample of 1000 rather than full 10.000
* What are the transformations we apply to the data?

1. Which do we put into dummy’s?
2. How do we scale the others?
3. What do we do with variables where some variables are missing?

**Method** (Ruben & Ruoying)

* **Describe in 2-3 sentences what a support vector machine is**

1. When is it appropriate to use it, and why is it better?
2. What is the general shape of the loss function?

* Why do use Huber errors?

1. Huber can approximate linear and quadratic if that’s better
2. But has flexibility (Trade-offs in terms of punishing outliers, smooth function)

* What is the adjusted rand index? (

1. Formula describing it (see: https://en.wikipedia.org/wiki/Rand\_index#Adjusted\_Rand\_index)
2. Why is it better than misclassification or F1 metric? Given sample is biased with lots of -1, index is created for avoiding random success

* Describe linear and non-homogeneous polynomial kernel
* Mention how we split it in train and test
* How to read the coefficient in SVM (discuss with Ruben)

**Results** (Floris & Ruoying)

* Briefly comment on how it yields the same results as svmmaj (plots in appendix)
* 2-3 sentences, graph in appendix to show similarity of results
* Show plots which hyper parameters (lambda, k\_huber) we ended up with
* 3D plot
* Then show result of applying these hyper parameters to a training and test set (70%/30% split on the sample 1000)
* Confusion matrix
* Adjusted Rand Index
* Plot to show which ones were wrong
* Then show what the effect of adding kernels is(we should discuss what we want to show here)
* For the model that works best, show a table with most important variables in terms of weight (+interpretation)