# Project plan

The project will consist of 3 models and a simulator. The simulator will be a basic heuristic used to benchmark all the other models and allow for comparison. The original data will also be tested in the simulator to be able to compare the historical (human-made) relocation moves.

The other 3 models are the following:

* Behavioural cloning: This model is the simplest as it only takes a current state and predicts an action based on what was done before. It does not attempt to improve the action, as it does not know how well it performed.
* Online reinforcement: The idea of building this model is to both have a benchmark and make the development of the offline reinforcement model easier. Given that an online environment would be to relocate car in reality, this model will use the simulator (and therefore rely on the simulator quality).
* Offline reinforcement: This model trains from the historical data (or possibly generated data from the simulator to make the comparison fairer) and it should be fairly similar in structure to the online reinforcement model.

Once all models are developed and perform as expected, the simulator will be used to compare the performance of all of them and make conclusions based on this. On the next page, a fortnight Gantt chart is attached with my estimated times for each stage.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Activity | 04-okt | 18-okt | 01-nov | 15-nov | 29-nov | 13-dec | 27-dec | 10-jan | 24-jan | 07-feb | 21-feb | 04-mar |
| Data exploration |  |  |  |  |  |  |  |  |  |  |  |  |
| Behaviour cloning model |  |  |  |  |  |  |  |  |  |  |  |  |
| Basic simulator development |  |  |  |  |  |  |  |  |  |  |  |  |
| Online reinforcement model |  |  |  |  |  |  |  |  |  |  |  |  |
| Offline reinforcement model |  |  |  |  |  |  |  |  |  |  |  |  |
| Writing thesis |  |  |  |  |  |  |  |  |  |  |  |  |