# Model Tuning

## Lecture

|  |  |
| --- | --- |
| Key messages | Questions from students |
| Regularisation is used to combat overfitting  **Ridge** gently pushes features to zero  **Lasso** harshly punishes unimportant parameters  Elastic net allows you to use both ridge and lasso for regularisation  P value is the probability that the co-efficient is not-statistically significant  Machine learning doesn’t necessarily have to remove features, as you can use regularisation to get rid of poor performing features- thus can adopt more of a brute force method  Note: You can use n=jobs to speed things up with parallelisation  Support Vector Machines are a robust prediction method, and have a good balance between results and compute power. Downside is that it shouldn’t be used for large data sets. It doesn’t provide probability outputs either.  The Hinge Loss function is applied to support vector machines to punish results that are applied to the wrong side of the margin. This is controlled with the hyper parameter C.  SVM.  Uses the concept of “Similarity”  Linear Kernal- Separates based on distance to hyperplane  Polynomial transforms the data so that it can be separated  RBF calculates the distance between two points, with points exponentially far away as a different class. |  |

## 1. Workflow

|  |  |
| --- | --- |
| Key messages | Questions from students |
|  |  |

## 2. Regularisation

|  |  |
| --- | --- |
| Key messages | Questions from students |
|  |  |

## 3. SVM

|  |  |
| --- | --- |
| Key messages | Questions from students |
|  |  |

## 4. Titanic

|  |  |
| --- | --- |
| Key messages | Questions from students |
|  |  |

## 5. Hand made SVM

|  |  |
| --- | --- |
| Key messages | Questions from students |
| into account for any of these functions |  |

## Recap

|  |  |
| --- | --- |
| Key messages | Questions from students |
|  |  |