Final reflection (mind map in final page)

1. Summary:

Things we learn this quarter:

Graphical user interface, text, application, chat or text message

Description automatically generated

1. Concepts:

**Models:**

Linear regression: use linear to fit data with Predictor:

Text

Description automatically generated

Regularized approach: ridge and Lasso and Elastic Net

Ridge Predictor:

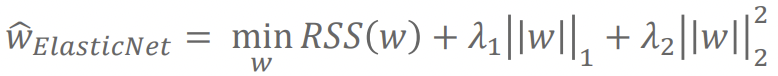
A picture containing text

Description automatically generated

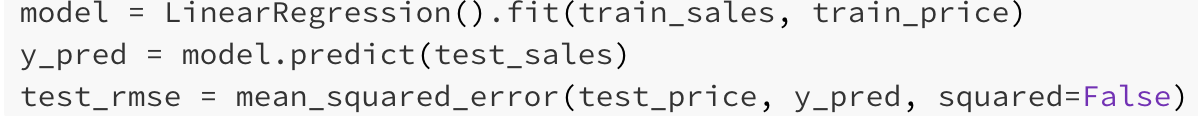
Lasso Predictor (have 0-coef variable which can be discarded:

Logo

Description automatically generated with low confidence

Elastic Net: 

Linear in code:



Ridge in code:

Text

Description automatically generated with medium confidence

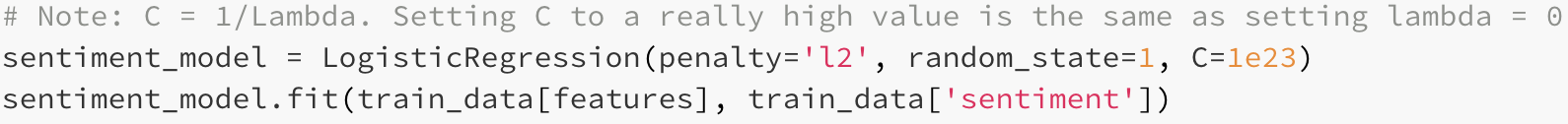
Lasso in code:

Text

Description automatically generated

Logistic regression (sentiment analysis):

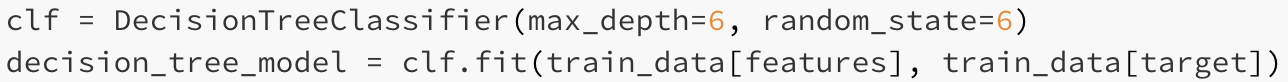
Give weights to words



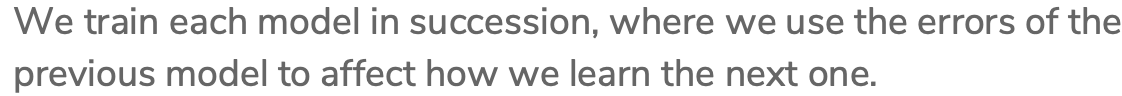
Non-linear model: decision tree

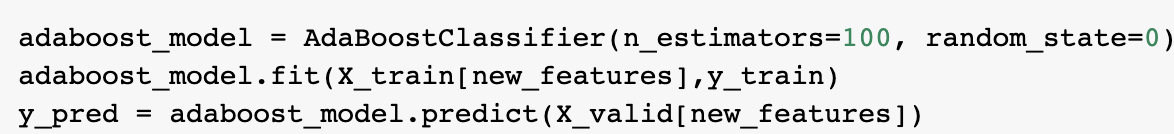
For categorical variable, we need to use get dummies method.





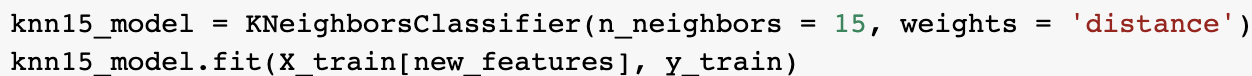
Adaboost:





Nearest neighbors, clustering: knn

Use majority classifier to judge the property of a word by seeing the nearest k words.



Recommendation system:



A picture containing text

Description automatically generated



Deep learning (use of cnn net):

Text

Description automatically generated

Algorithm:

Gradient accent:

A picture containing diagram

Description automatically generated

Chart, surface chart

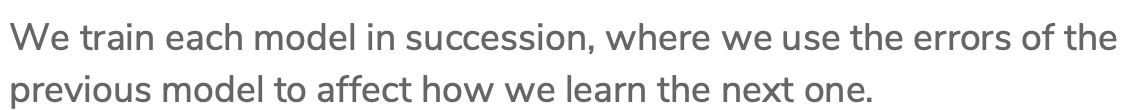
Description automatically generated Gradient decent is in opposite way.

Random forest (bagging):

Graphical user interface, text, application, email

Description automatically generated

Adaboost (boosting):



Text

Description automatically generated with medium confidence

k-means: clustering (unsupervised)

Graphical user interface, text, application, email

Description automatically generated

Point estimation, MLE (maximize likelihood estimate)

Diagram, schematic

Description automatically generated

Bias-variance trade off:



High complexity has high variance

Graphical user interface, application

Description automatically generated

I don’t know how to add an extra column to the left, so I wrote below:

Regression: use linear to approximate data

Classification: judge if a sentence is good or bad

Recommendation system: recommend by known user habits

Deep learning: identify the object by seeing the feature of what most of such object looks like.

Clustering: cluster similar data.

1. Uncertainty:

I think there is no so confuse concept during this class maybe. Thanks for the whole quarter. I am very appreciate the required learning reflection.