IN5550: Neural Methods in Natural Language Processing
Sub-lecture 2.1
Supervised Machine Learning

Andrey Kutuzov

University of Oslo

31 January 2023

Contents

I am Andrey Kutuzov

I will do some of the lectures, covering the following topics:

I am Andrey Kutuzov

I will do some of the lectures, covering the following topics:

- a review of supervised learning (introducing notation)

Introduction

I am Andrey Kutuzov

I will do some of the lectures, covering the following topics:

- a review of supervised learning (introducing notation)
- linear classifiers and simple feed-forward neural networks

Introduction

I am Andrey Kutuzov

I will do some of the lectures, covering the following topics:

- a review of supervised learning (introducing notation)
- linear classifiers and simple feed-forward neural networks
- multi-layer neural networks and their training

I am Andrey Kutuzov

I will do some of the lectures, covering the following topics:

- a review of supervised learning (introducing notation)
- linear classifiers and simple feed-forward neural networks
- multi-layer neural networks and their training
- language modeling and distributed word embeddings
- recurrent neural networks (RNNs) and contextualized embeddings

I am Andrey Kutuzov

I will do some of the lectures, covering the following topics:

- a review of supervised learning (introducing notation)
- linear classifiers and simple feed-forward neural networks
- multi-layer neural networks and their training
- language modeling and distributed word embeddings
- recurrent neural networks (RNNs) and contextualized embeddings

I am also partially responsible for the obligatory assignments:

1. Bag of Words Document Classification
2. ...
3. ...

Technicalities

- Familiarize yourself with the course infrastructure.

Introduction

Technicalities

- Familiarize yourself with the course infrastructure.
- Check the course page for messages.

Technicalities

- Familiarize yourself with the course infrastructure.
- Check the course page for messages.
- Test whether you can access `https://github.uio.no/in5550/2023`
  - make sure to update your UiO GitHub profile with your photo, and star the course repository :-)

Technicalities

- Familiarize yourself with the course infrastructure.
- Check the course page for messages.
- Test whether you can access `https://github.uio.no/in5550/2023`
  - make sure to update your UiO GitHub profile with your photo, and star the course repository :-)
- Most of machine learning revolves around linear algebra.

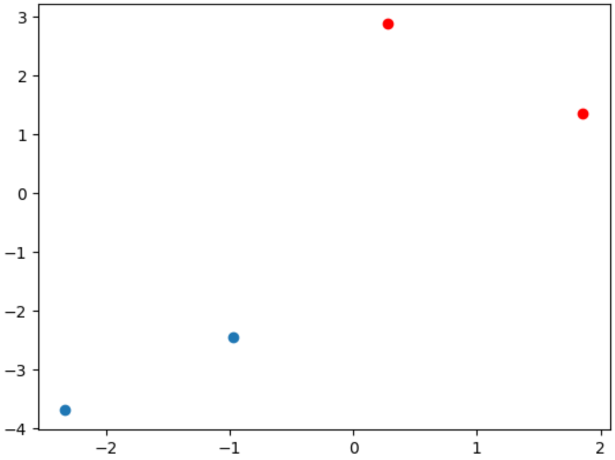
Technicalities

- Familiarize yourself with the course infrastructure.
- Check the course page for messages.
- Test whether you can access `https://github.uio.no/in5550/2023`
  - make sure to update your UiO GitHub profile with your photo, and star the course repository :-)
- Most of machine learning revolves around linear algebra.
- We created a *LinAlg* cheat sheet for this course.

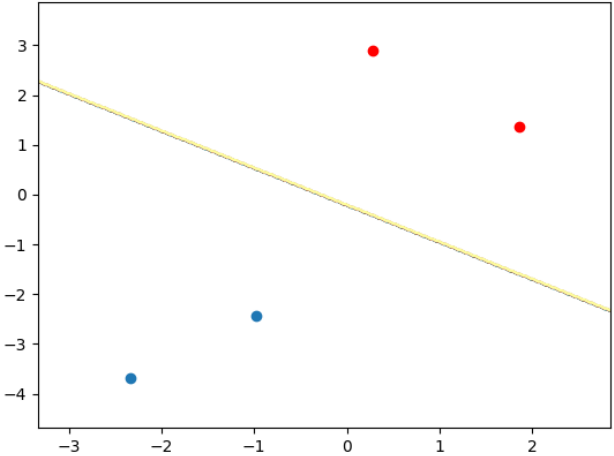
Technicalities

- Familiarize yourself with the course infrastructure.
- Check the course page for messages.
- Test whether you can access `https://github.uio.no/in5550/2023`
  - make sure to update your UiO GitHub profile with your photo, and star the course repository :-)
- Most of machine learning revolves around linear algebra.
- We created a *LinAlg* cheat sheet for this course.
  - Linked from the course page, adapted for the notation of [Goldberg, 2017].

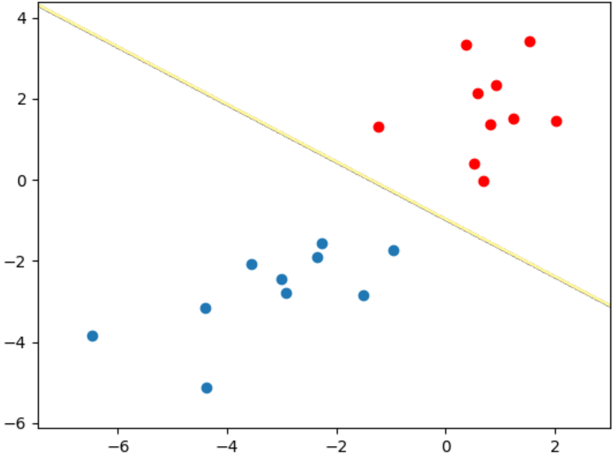
Contents


3
2
1
0
−1
−2
−3
−4
−2
−1
0
1
2

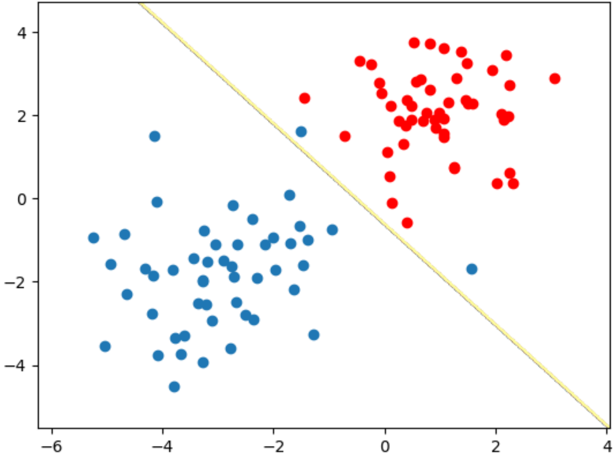
Basics of supervised machine learning



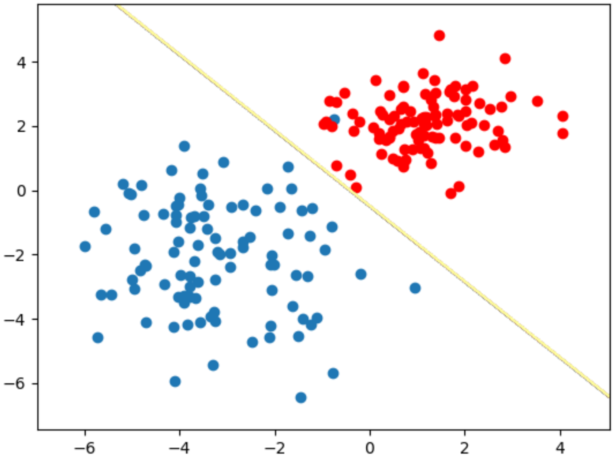
Basics of supervised machine learning



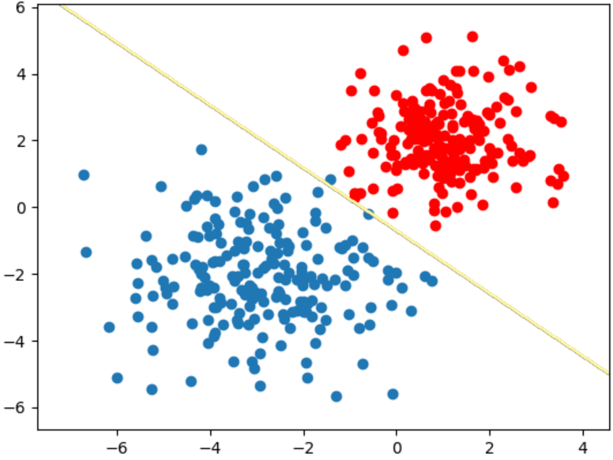
Basics of supervised machine learning



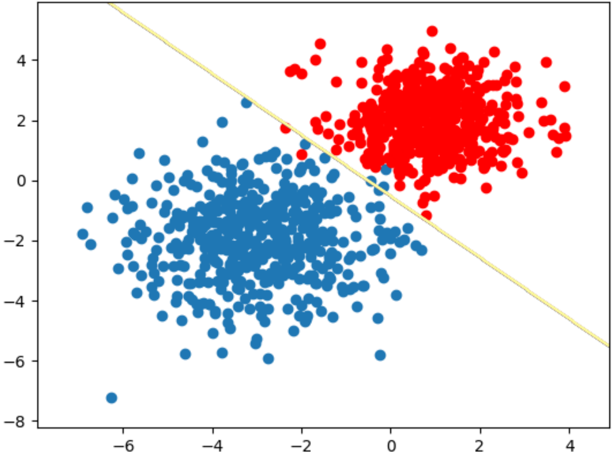
Basics of supervised machine learning



Basics of supervised machine learning



Basics of supervised machine learning



- Input 1: a training set of n training instances $x_{1:n} = x_1, x_2, \ldots x_n$

Basics of supervised machine learning

- Input 1: a training set of n training instances $x_{1:n} = x_1, x_2, \ldots x_n$
  - for example, e-mail messages.

Basics of supervised machine learning

- Input 1: a training set of n training instances $x_{1:n} = x_1, x_2, \ldots x_n$
  - for example, e-mail messages.
- Input 2: corresponding 'gold' labels for these instances $y_{1:n} = y_1, y_2, \ldots y_n$

Basics of supervised machine learning

- Input 1: a training set of n training instances $x_{1:n} = x_1, x_2, \ldots x_n$
  - for example, e-mail messages.
- Input 2: corresponding 'gold' labels for these instances $y_{1:n} = y_1, y_2, \ldots y_n$
  - for example, whether the message is spam (1) or not (0).

Basics of supervised machine learning

- Input 1: a training set of n training instances $x_{1:n} = x_1, x_2, \ldots x_n$
  - for example, e-mail messages.
- Input 2: corresponding 'gold' labels for these instances $y_{1:n} = y_1, y_2, \ldots y_n$
  - for example, whether the message is spam (1) or not (0).
- The trained models allow to make label predictions for unseen instances.

Basics of supervised machine learning

- Input 1: a training set of n training instances $x_{1:n} = x_1, x_2, \ldots x_n$
  - for example, e-mail messages.
- Input 2: corresponding 'gold' labels for these instances $y_{1:n} = y_1, y_2, \ldots y_n$
  - for example, whether the message is spam (1) or not (0).
- The trained models allow to make label predictions for unseen instances.
- Generally: some program for mapping instances to labels.

Recap on data split

- Recall: we want the model to make good predictions for unseen data.

Basics of supervised machine learning

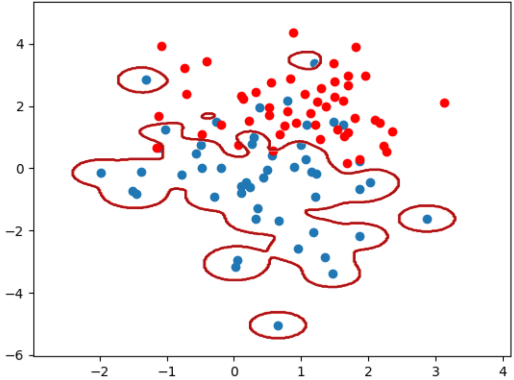
Recap on data split

- Recall: we want the model to make good predictions for unseen data.
- It should not overfit to the seen data.

Basics of supervised machine learning

Recap on data split

- Recall: we want the model to make good predictions for unseen data.
- It should not overfit to the seen data.



Basics of supervised machine learning

Remember: we want models that generalize

- Thus, the datasets are usually split into:
  1. train data;

Basics of supervised machine learning

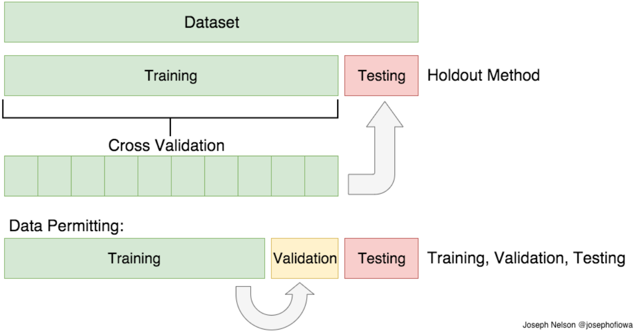
Remember: we want models that generalize

- Thus, the datasets are usually split into:
  1. train data;
  2. validation/development data (optional);

Basics of supervised machine learning

Remember: we want models that generalize

- Thus, the datasets are usually split into:
  1. train data;
  2. validation/development data (optional);
  3. test/held-out data.



Basics of supervised machine learning

- We want to find a function which makes good, generalizable predictions for our task.

Basics of supervised machine learning

- We want to find a function which makes good, generalizable predictions for our task.
- Searching among all possible functions is unfeasible.

Basics of supervised machine learning

- We want to find a function which makes good, generalizable predictions for our task.
- Searching among all possible functions is unfeasible.
- To cope with that, we choose an inductive bias...

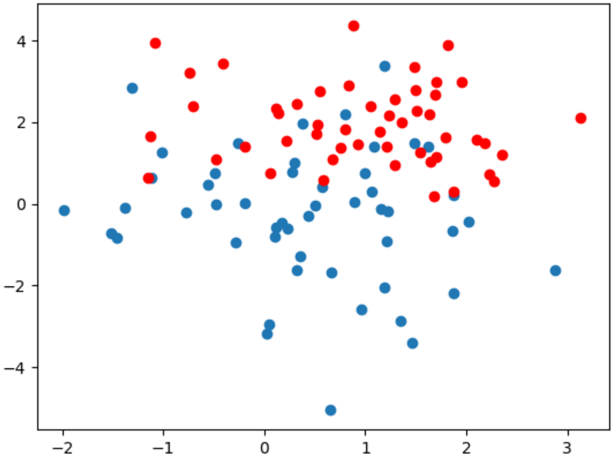
Basics of supervised machine learning

- We want to find a function which makes good, generalizable predictions for our task.
- Searching among all possible functions is unfeasible.
- To cope with that, we choose an inductive bias...
- ...and set some hypothesis class...

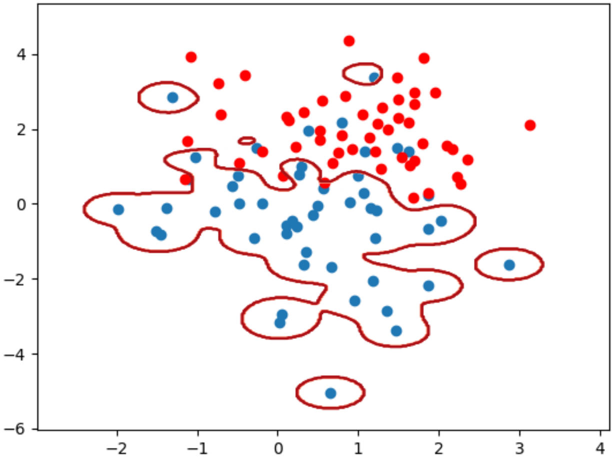
- We want to find a function which makes good, generalizable predictions for our task.
- Searching among all possible functions is unfeasible.
- To cope with that, we choose an inductive bias...
- ...and set some hypothesis class...
- ...to search only within this class.

What do you think a good model would look like for this data?

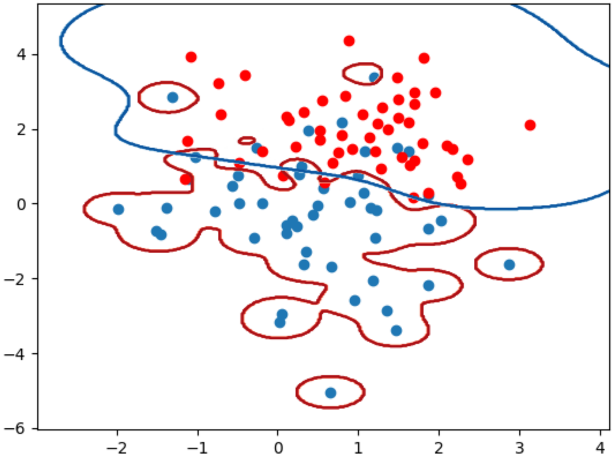
What do you think a good model would look like for this data?



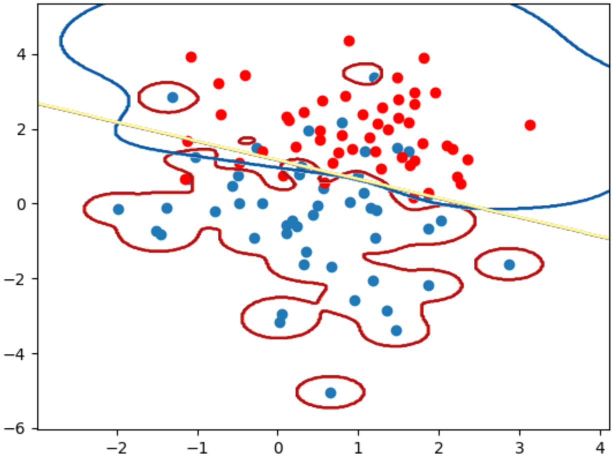
What do you think a good model would look like for this data?



What do you think a good model would look like for this data?



What do you think a good model would look like for this data?



Linear functions: a popular hypothesis class

(move on to sub-lecture 2.2)

References I

Goldberg, Y. (2017).
Neural network methods for natural language processing.
Synthesis Lectures on Human Language Technologies, 10(1):1–309.