Tutorial 9 - Class Imbalance & K-means Clustering

Victoria Ajila, MASc Computer Engineering
Carleton University

Monday 22nd November, 2021

Disclaimer: Recorded Tutorials will be Posted

Privacy Preservation:

- Ask questions in the chat[1]
- Keep video off

Note: If the above *hinders your ability to learn* $\wedge$ *violates your privacy*, please let me/Dr. Green know ASAP and video will be post-processed accordingly.

[1] I encourage unmuted/voice-based questions at any time, but know that this isn't explicitly privacy-preserving

ML Weekly

Recent news events from the ML community

ML Weekly

1. **(ML)** "Deepfaking the Mind" Could Improve Brain-Computer Interfaces for People with Disabilities

ML Weekly

1. **(ML)** "Deepfaking the Mind" Could Improve Brain-Computer Interfaces for People with Disabilities
2. **(AI)** Giving robots social skills

ML Weekly

1. **(ML)** "Deepfaking the Mind" Could Improve Brain-Computer Interfaces for People with Disabilities
2. **(AI)** Giving robots social skills
3. **(NLP)** Toward speech recognition for uncommon spoken languages

Tutorial Overview

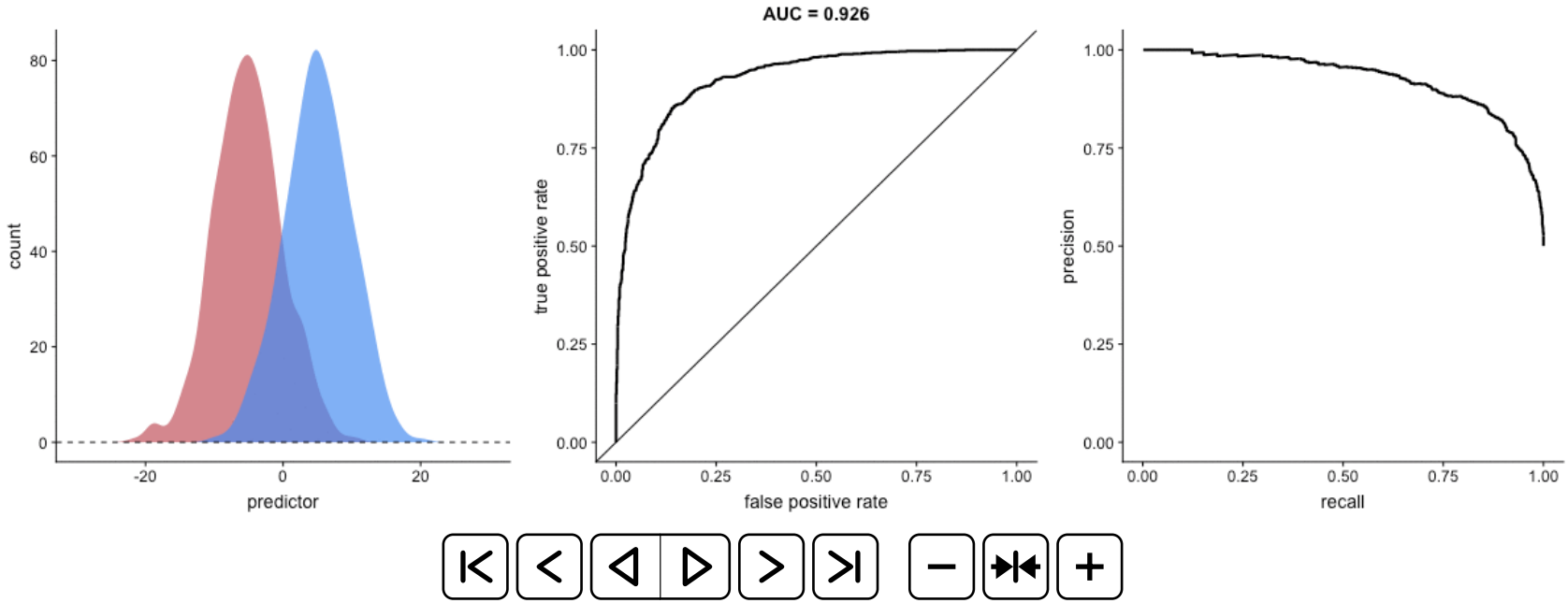
We will cover two main concepts in the notebook today:

1. Class Imbalance and K-means clustering

Tutorial Intuition

Building an Intuition for the Concepts of this Tutorial

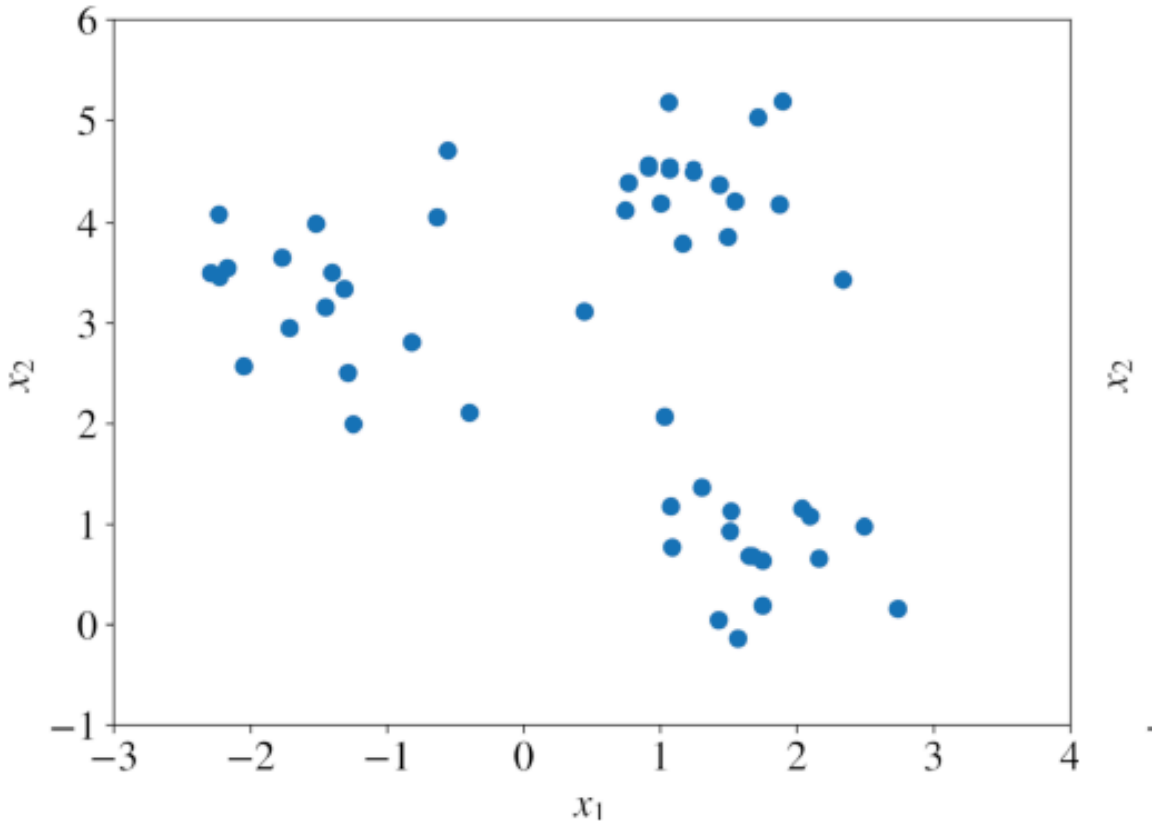
Animation: Variation in Class Imbalance



K-means Clustering Intuition

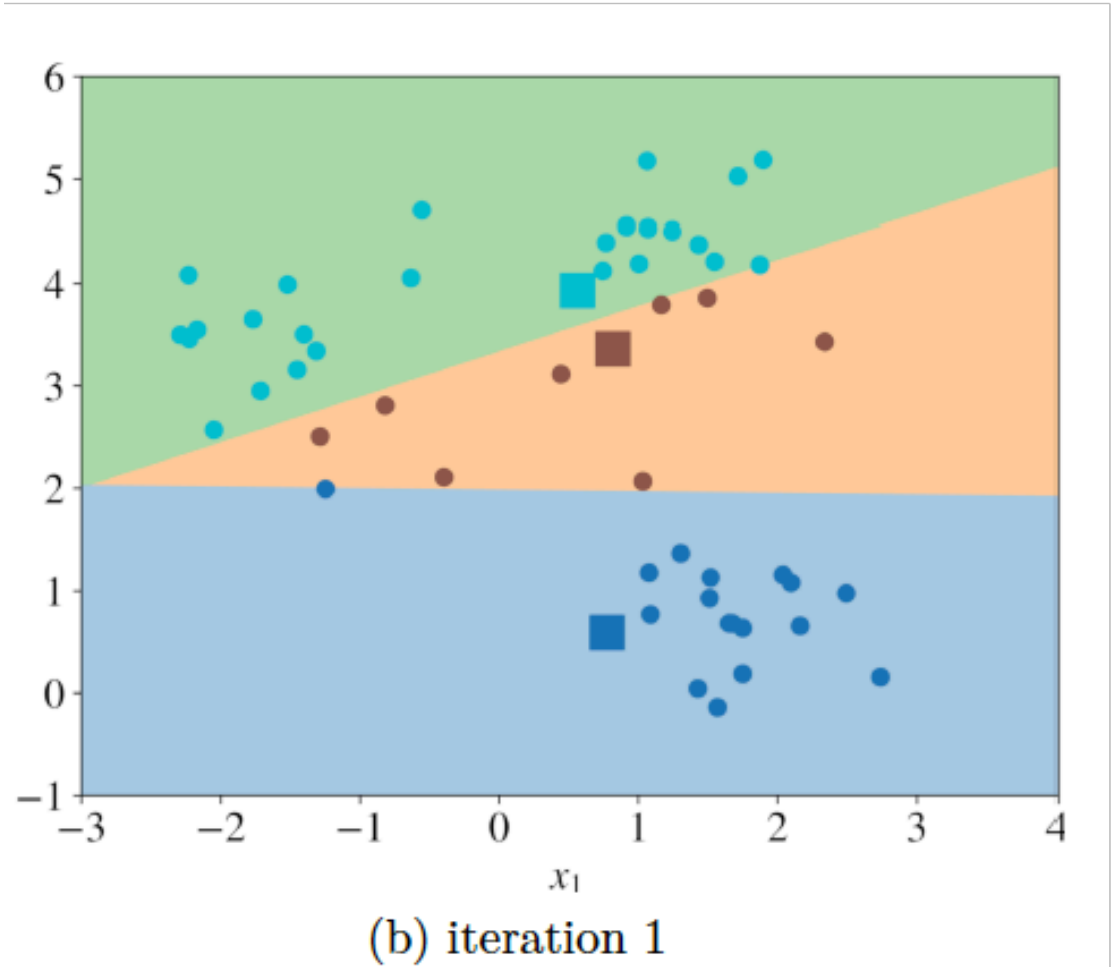
1. Choose k (the number of clusters)
2. Randomly place k feature vectors, called centroids, to the feature space
3. Compute the distance from each example x to each centroid c using some metric, (ex. Euclidean distance)
4. Then we assign the closest centroid to each example
5. For each centroid, we calculate the average feature vector of the examples labeled with it. The average feature vectors become the new locations of the centroids.
6. Repeat 3-5 until the assignments don't change after the centroid locations were recomputed

K-means Clustering Intuition



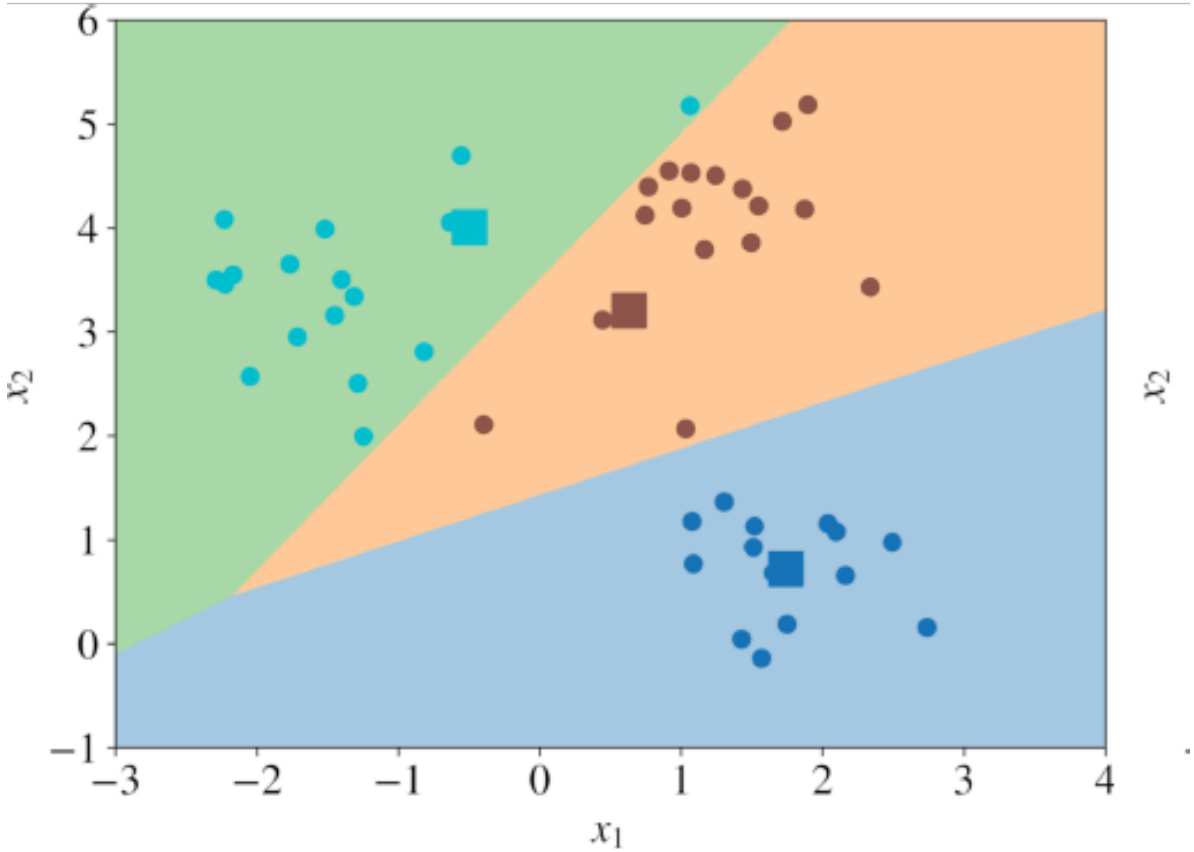
(a) original data

K-means Clustering Intuition



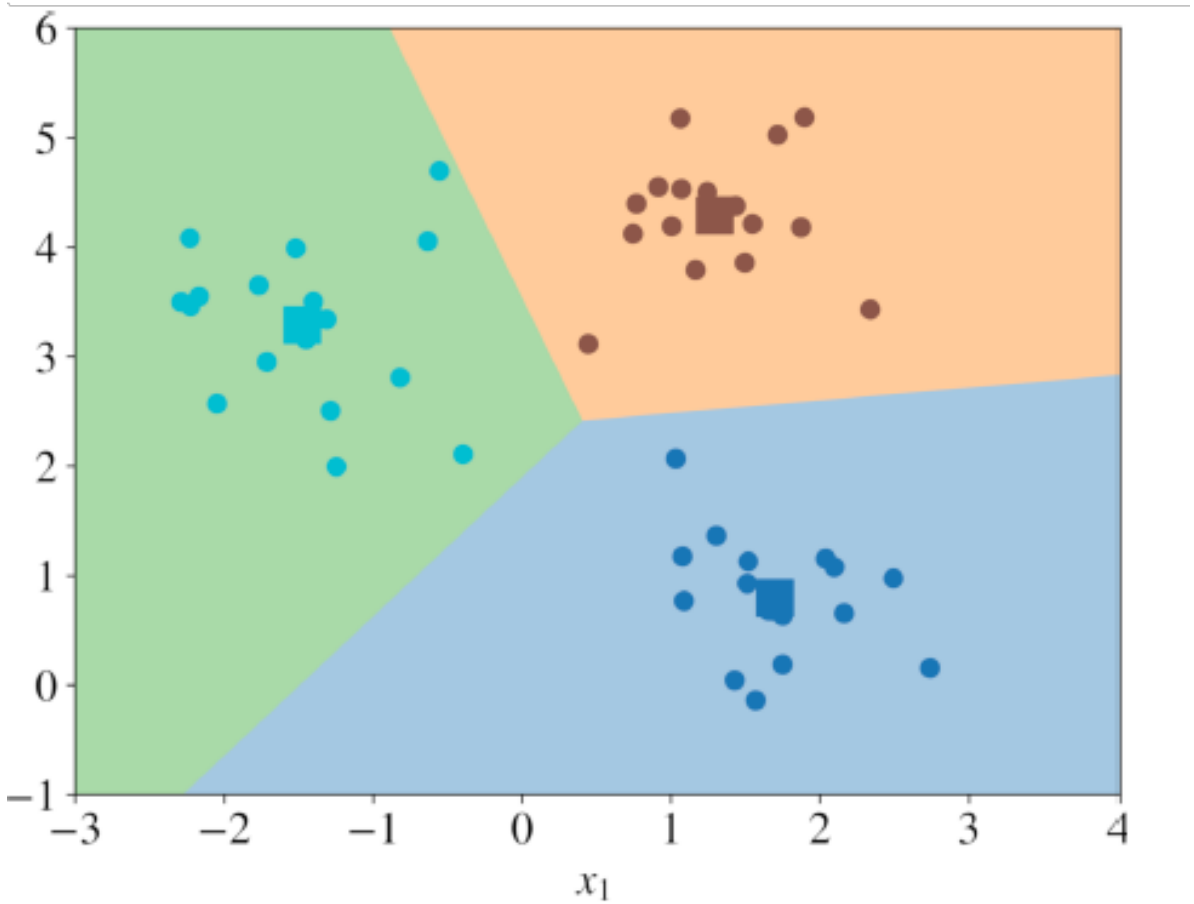
(b) iteration 1

K-means Clustering Intuition



(c) iteration 3

K-means Clustering Intuition



(d) iteration 5

Into the Notebooks we Go...

We will cover two notebooks today!

1. Tutorial 9 - Class Imbalance and K-means clustering

Tutorial 9 - Class Imbalance & K-means Clustering

Victoria Ajila, MASc Computer Engineering
Carleton University

Monday 22nd November, 2021