Holistic AI × The Alan Turing Institute

Bias in Clustering Systems Part I

Sachin Beepath

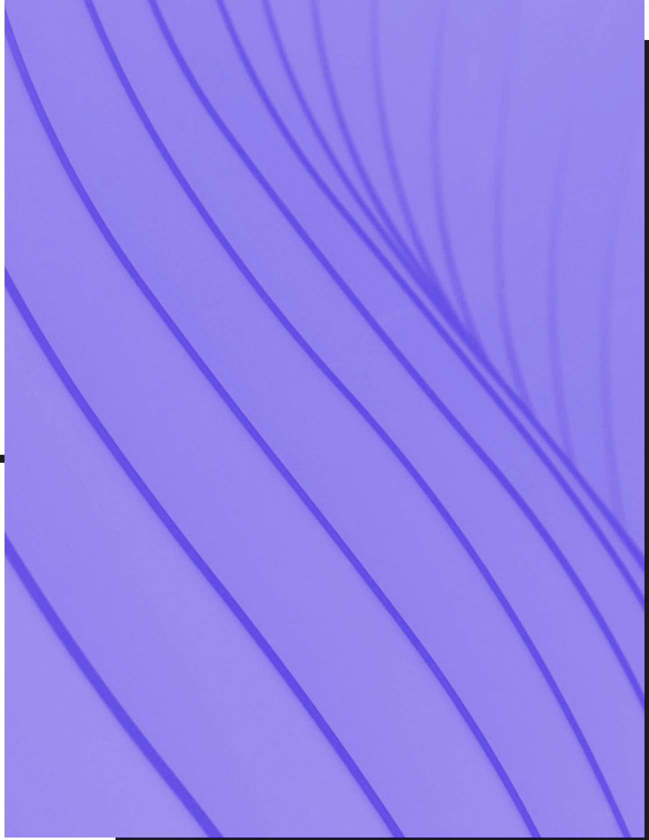
Assurance Officer, Holistic AI

Contents

Contents

Introduction to Clustering

1. Introduce Clustering as a form of AI.
2. Provide real world examples of Clustering.
3. Motivate the importance of fairness in Clustering.



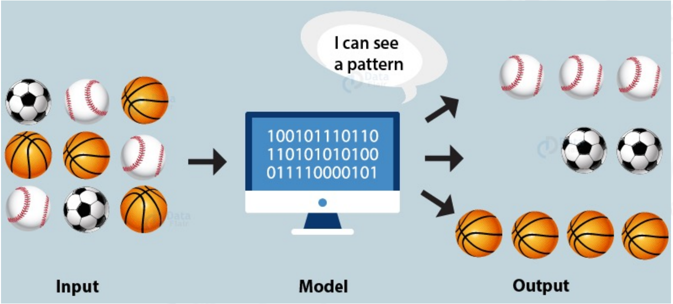
What is Clustering?

– Unsupervised learning method that groups objects into ‘clusters’.

– Objects in each cluster are more ‘similar’ to each other than to objects in another.

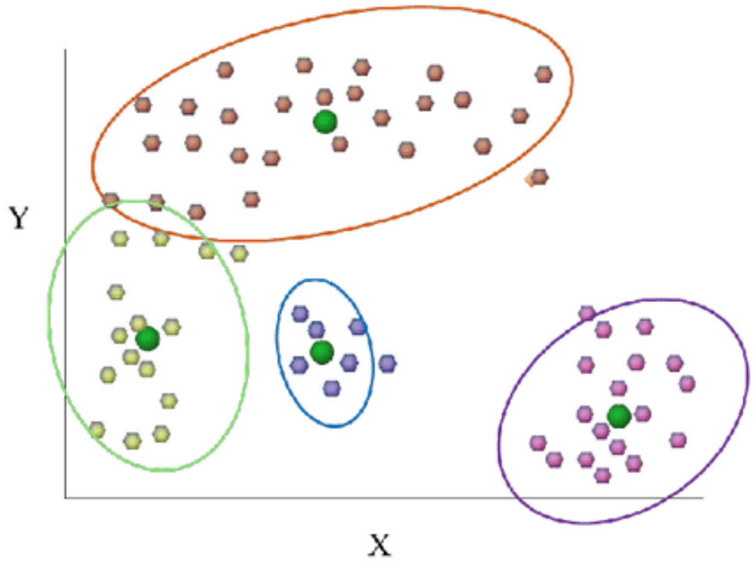
– Gain insight into unlabelled datasets.

– Playlist Curation.



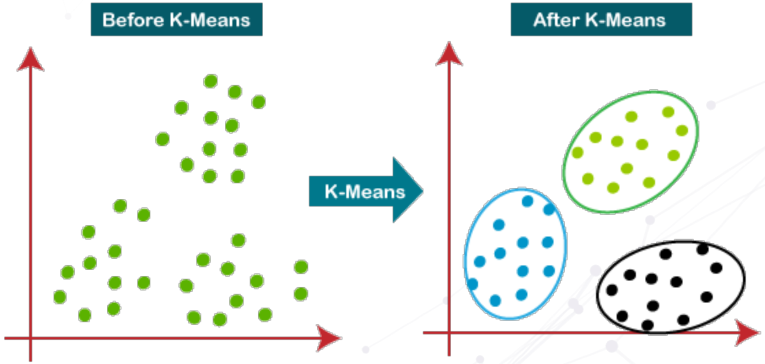
Centre-Based Clustering

- Split the data into k groups with a *centroid* to represent each cluster.
- Points are assigned to cluster with the most ‘similar’ centroid.
- Distance is used as a ‘dissimilarity’ metric.
- Aims to assign objects to the group with the closest centroid.
- K-means, K-medoids, mean-shift.



K-Means Clustering Algorithm

- Randomly choose k points as centroids.
- Measure the distance between each point and each centroids.
- Assign the cluster with the smallest centroid distance.
- Update each centroid to be the mean of the cluster.
- Repeat until the cluster assignments and centroids cease to change.



K-Means Clustering Algorithm

– Given k cluster centroids and n data points, the solution for k-means is:

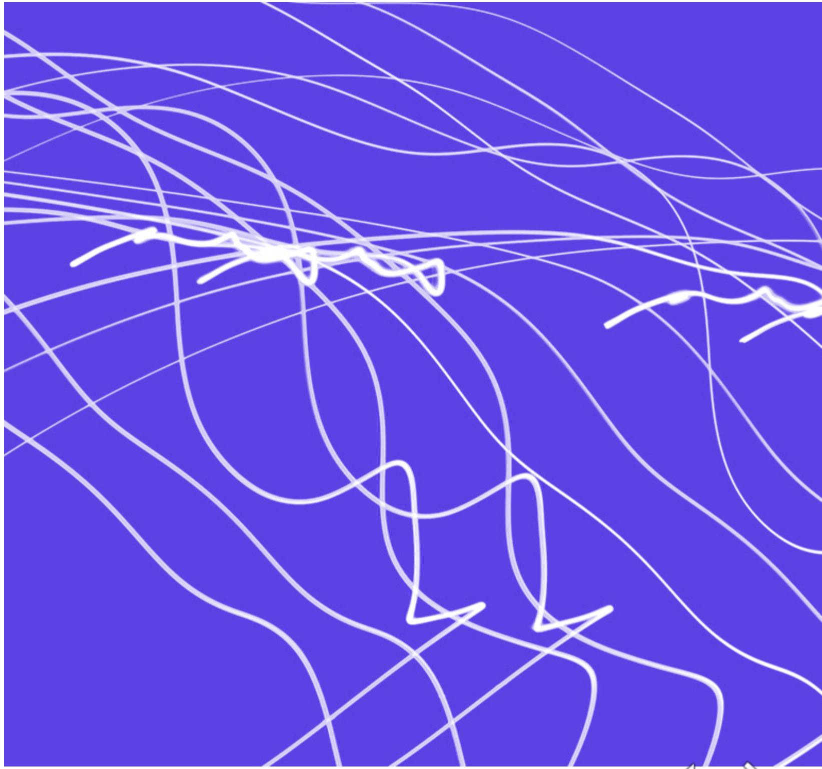
$$C \in \{c_1, c_2, \ldots, c_k\},\ c_i \in \Re^d$$

$$a \in \{a_1, a_2, \ldots, a_n\},\ a_i \in \{1, 2, \ldots, k\}$$

$$\hat{C}, \hat{a} = \underset{C,a}{\arg\min} \sum_{i}^{n} \|x_i - c_{a_i}\|^2$$

K-Means Clustering Algorithm

Pros	Cons
Easy to implement.	Not guaranteed to converge to optimal solution.
Easily interpreted.	Strongly depends on initial choice of centroids.
Guaranteed to converge to *a* solution.	Sensitive to outliers.



Example of K-Means Clustering

- Drafting players for a fantasy premier league team
- Choose similar players.
- Using the 2021-2022 Premier League season data.

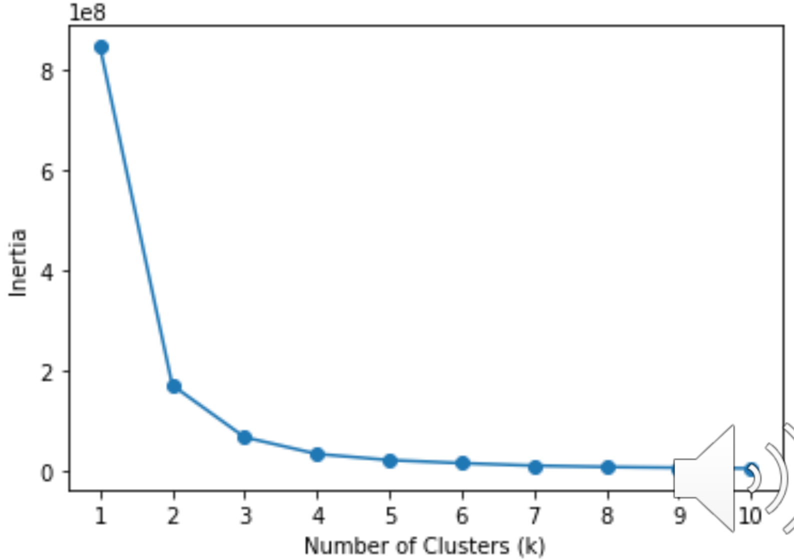
(https://github.com/vaastav/Fantasy-Premier-League/)

	first_name	second_name	goals_scored	assists	total_points	minutes	goals_conceded	ict_index	clean_sheets	red_cards	yellow_cards	element_type
0	Bernd	Leno	0	0	10	360	9	8.5	1	0	0	GK
1	Rúnar Alex	Rúnarsson	0	0	0	0	0	0.0	0	0	0	GK
2	Willian	Borges Da Silva	0	0	0	0	0	0.0	0	0	0	MID
3	Pierre-Emerick	Aubameyang	4	1	44	1036	16	92.9	6	0	3	FWD
4	Cédric	Soares	1	1	48	1481	27	63.0	3	0	3	DEF

Drafting a Fantasy Team with Clustering

- Train several clustering model using all numerical input features.
- Choose optimal number of clusters using elbow method.

```python
# create a KMeans instance with k clusters
model = KMeans(n_clusters = k)
# fit model to samples
model.fit(data)
```



Drafting a Fantasy Team with Clustering

- Look at which players to draft based on the clusters desirable players get assigned to and choosing players from the same clusters.
- If we take the cluster that Harry Kane is in, we can see all the players that are 'similar' to him.

	first_name	second_name	goals_scored	assists	total_points	minutes	goals_conceded	ict_index	clean_sheets	red_cards	yellow_cards	element_type	cluster
19	Bukayo	Saka	11	9	179	2978	39	311.8	13	0	6	MID	3
20	Gabriel	Magalhães	5	0	146	3063	38	128.4	13	1	6	DEF	3
25	Ben	White	0	0	107	2880	35	92.5	13	0	3	DEF	3
31	Martin	Ødegaard	7	4	131	2782	39	220.3	11	0	4	MID	3
32	Aaron	Ramsdale	0	0	135	3060	39	68.0	12	0	1	GK	3

Part 1 Question 1

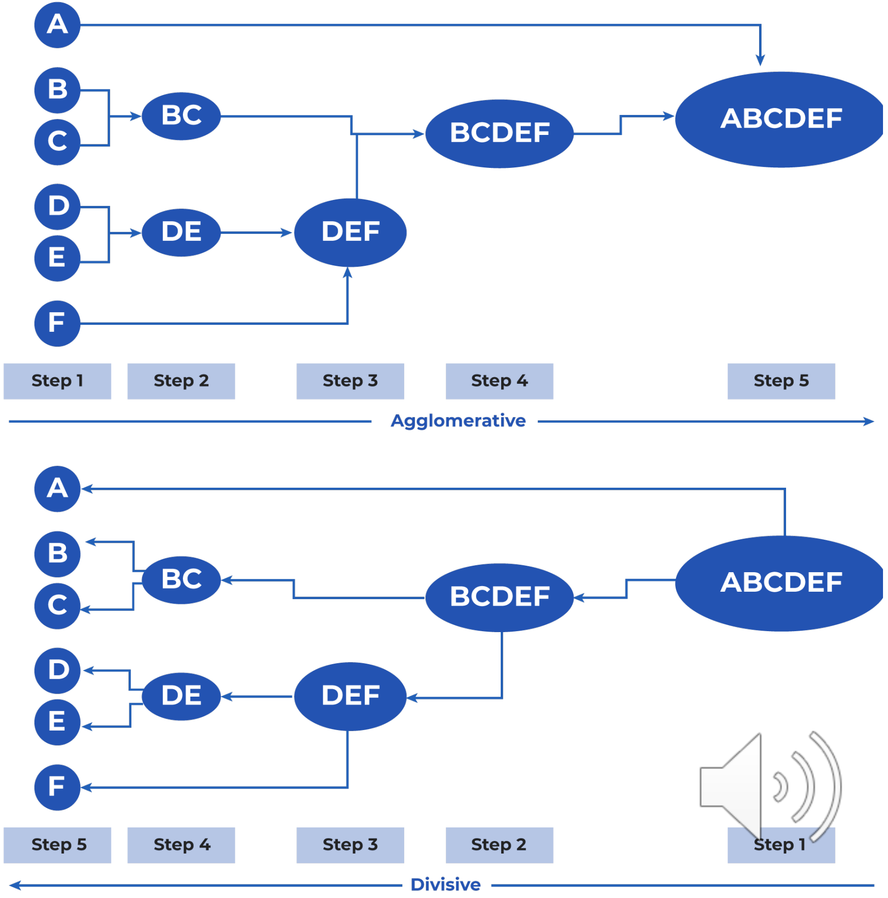
Given these 4 points and distances to cluster centres, how many points get assigned to to each cluster?

X	0.78	2.23	1.57	1.21
Y	-1.86	-2.67	-2.76	-2.55
Distance from first centroid	3.62	4.48	4.49	4.27
Distance from first centroid	4.16	5.15	5.11	4.86

Other forms of Clustering

Hierarchical

- Partitions data set into hierarchies.
- Agglomerative builds the tree from the bottom up.
- Divisive builds it from the top down.
- The root of the tree represents the entire data set while the leaves compromise singular samples.
- Removes the need to have a predefined number of clusters.
- Customer segmentation.



Other forms of Clustering

Mixture model

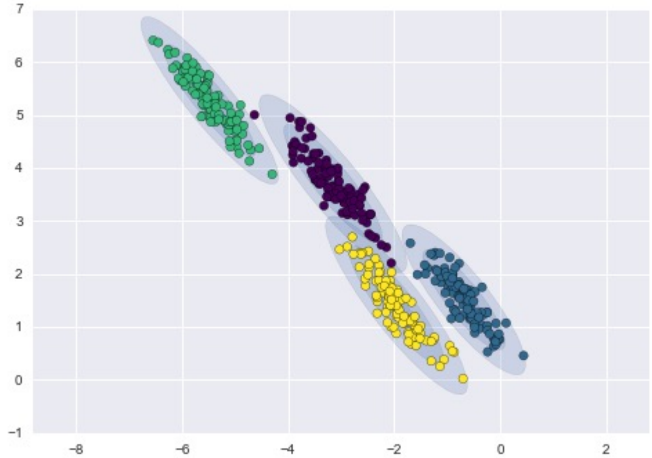
– Probabilistic clustering.

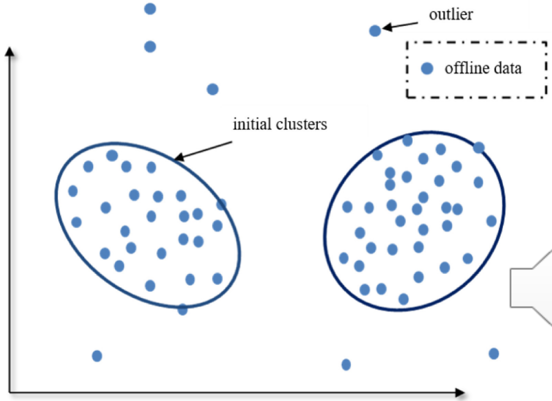
– Points are assigned to clusters with no fixed membership.

– Data points are assumed to come from some mixture of probability distributions.

– Distributions are assumed.

– GMM-EM for anomaly detection.





Other forms of Clustering

Spectral

- Construct a graph to encode similarity between all data points.
- The weight of the edge between nodes is a function of how similar they are to each other.
- Weighted graph represented matrix W where $w_{i,j}$ is similarity between point x_i and point x_j.
- Degree of the i-th node is defined as $d_i = \sum_j^n w_{ij}$, can construct the degree matrix D.
- The graph Lapacian Matrix is defined as: $L = D - W$.

Other forms of Clustering

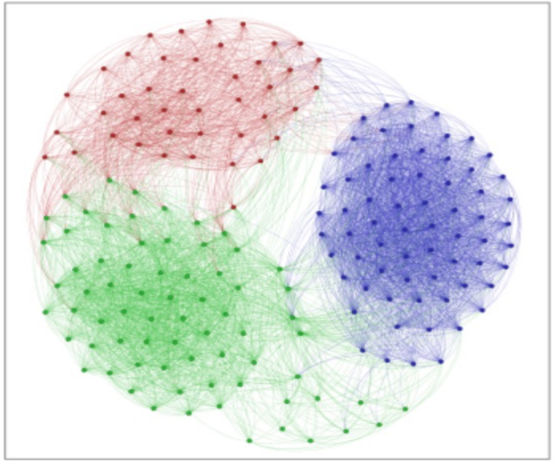
Spectral (cont.)

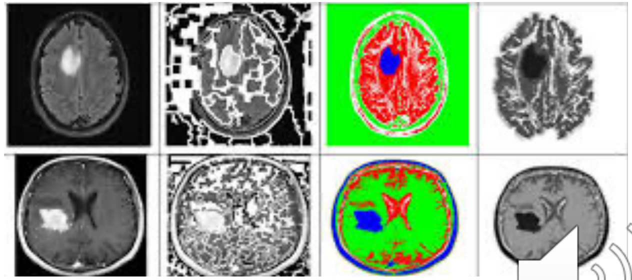
– The k smallest eigenvectors of L can be used to construct a lowFromer-dimensional matrix V using them as column vectors.

– From the new matrix we create a new set of data points using its row vectors: $y_1, \ldots, y_n$.

– Cluster new dataset using k-means and assign any point x_j from the old dataset to the same cluster its corresponding point in the new dataset y_j was assigned to.

– Image segmentation.





Part 1 Question 2

What type of clustering might be best suited if you would like to automatically group documents together?

A. Centre-Based
B. Hierarchical
C. Mixture Model
D. Spectral

Part 1 Question 2

What type of clustering might be best suited if you would like to automatically group documents together?

A. Centre-Based

B. Hierarchical

C. Mixture Model

D. Spectral

How Can Clustering Be Unfair?

Bank Loans

- Used to determine whether someone should get a loan based on how likely they are to default on it.
- Input features often contain proxies to protected attributes (income, education, marital status).
- Due to historical systemic discrimination, women and POC may have lower incomes.
- Married people typically have higher credit scores than single people.
- This clustering can result in rejection or higher interest rates.

How Can Clustering Be Unfair? (Cont.)

Job Shortlisting

- Group similar candidates together
- Choose who should move on to the next round
- Can reduce human bias but does not eliminate bias completely
- Features relevant to choosing a candidate may contain proxies or the protected attribute themselves

How Can Clustering Be Unfair? (Cont.)

Prisoner Recidivism

– The tendency for a prisoner to reoffend can be interpreted as a probability.

– Probability can be determined by a soft clustering algorithm (GMM-EM).

– A data point can be assigned a certain proportion of each cluster with clusters signifying the level of risk of re-offending .

– COMPAS.

Conclusion

– Introduced Clustering as a form of AI

– Seen the applications of Clustering and how it can be used in real world examples

– Explored how Clustering can be biased and the need for fairness

Contents