v-Means: A Visible and Explainable Clustering Algorithm

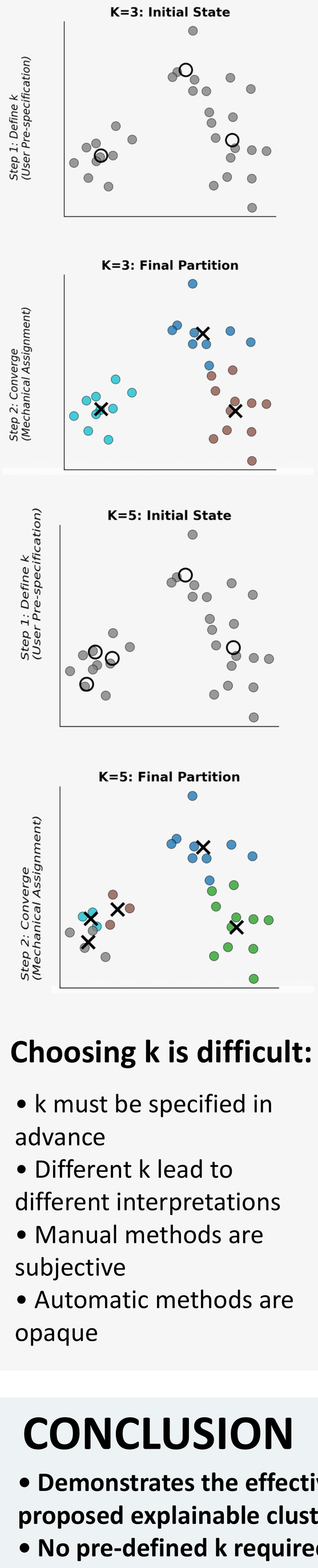
Liantian Shan and Robert S. Laramee
The School of Computer Science, The University of Nottingham, UK

PRASE 1:INITIALIZATION | PHASE 2:SEGMENIATION | PHASE 3: ORADIENT DETECTION | PHASE 4: GLUSTERING

PROBLEM:

How Many Clusters?

k-means is opaque

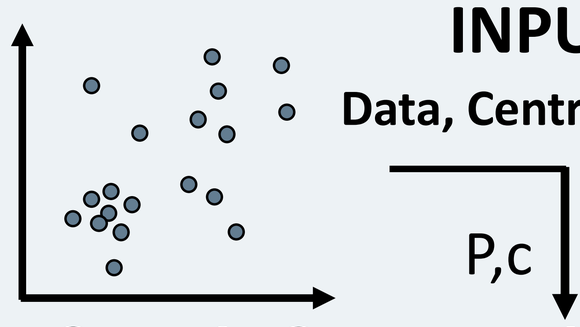


Choosing k is difficult:

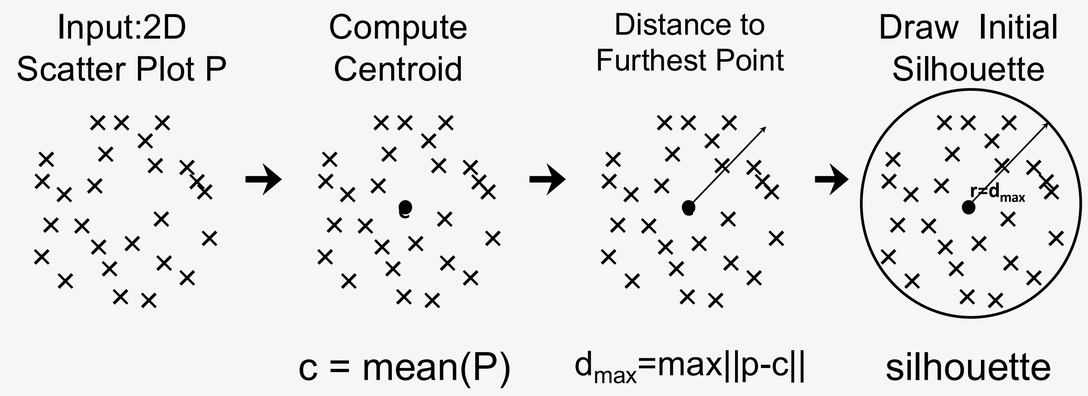
- k must be specified in advance
- Different k lead to different interpretations
- Manual methods are subjective
- Automatic methods are opaque

VISUAL EXPLANATION:

The Silhouette-based Recursive Pipeline



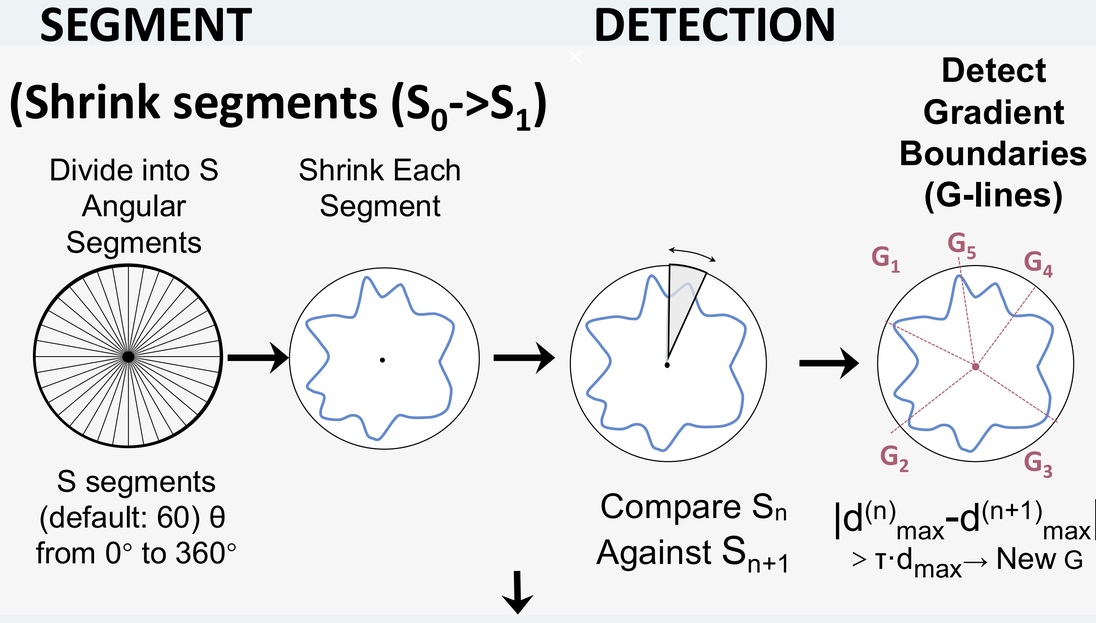
Step 1: Compute Initial Silhouette



STEP 2: SHRINK EACH SEGMENT

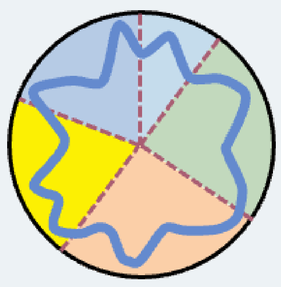
STEP3: GRADIENT DETECTION

(Shrink segments (S_0->S_1)



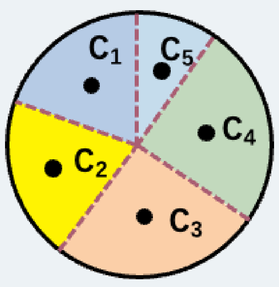
STEP 4: CLUSTERING

SPLIT DATA USING G-LINES



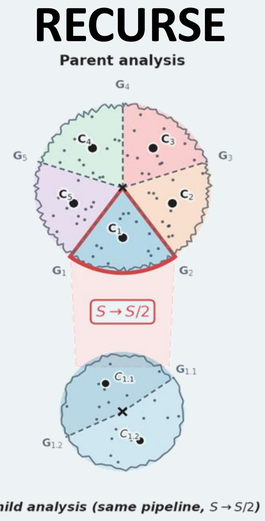
Split Data Using G-lines

COMPUTE NEW CENTROIDS



Output Cluster Centers C

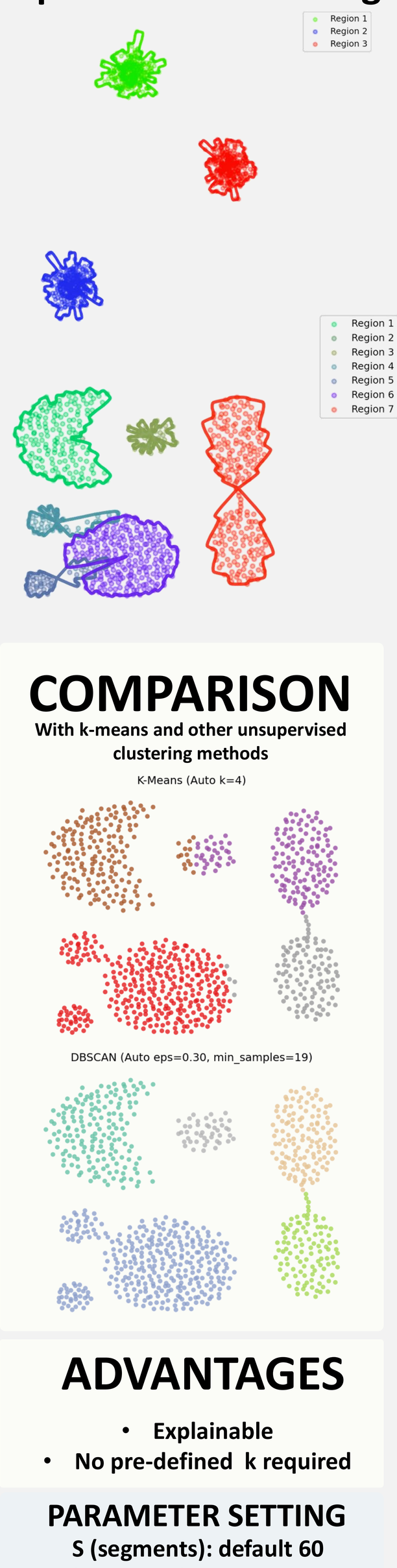
RECURSE



Recursion (S→S/2 for each recursive selected region)

RESULTS:

Automatic & Explainable Clustering



COMPARISON

With k-means and other unsupervised clustering methods

ADVANTAGES

- **Explainable**
- **No pre-defined k required**

PARAMETER SETTING

S (segments): default 60
(threshold): default 0.25

CONCLUSION

- **Demonstrates the effectiveness of the proposed explainable clustering method**
- **No pre-defined k required**
- **Provides a step-by-step visual explanation of cluster formation**

FUTURE WORK

- **A user study to evaluate explainability**
- **A hierarchical implementation of the algorithm**

CONTACT

contact
WhatsApp