

Clustering Ashkenazi Manuscripts

Seeking sub-clusters within the medieval Ashkenazi square script



Funded by the European Union



European Research Council

Mohammad Suliman
suliman@mail.tau.ac.il

Berat Kurar Barakat
berat@tauex.tau.ac.il

Sharva Gogawale
sharvag@mail.tau.ac.il

Daria Vasyutinsky Shapira
dariashap@tauex.tau.ac.il

Nachum Dershowitz
nachum@tau.ac.il

Ashkenazi square script clustering (ASC) dataset

German samples

French samples

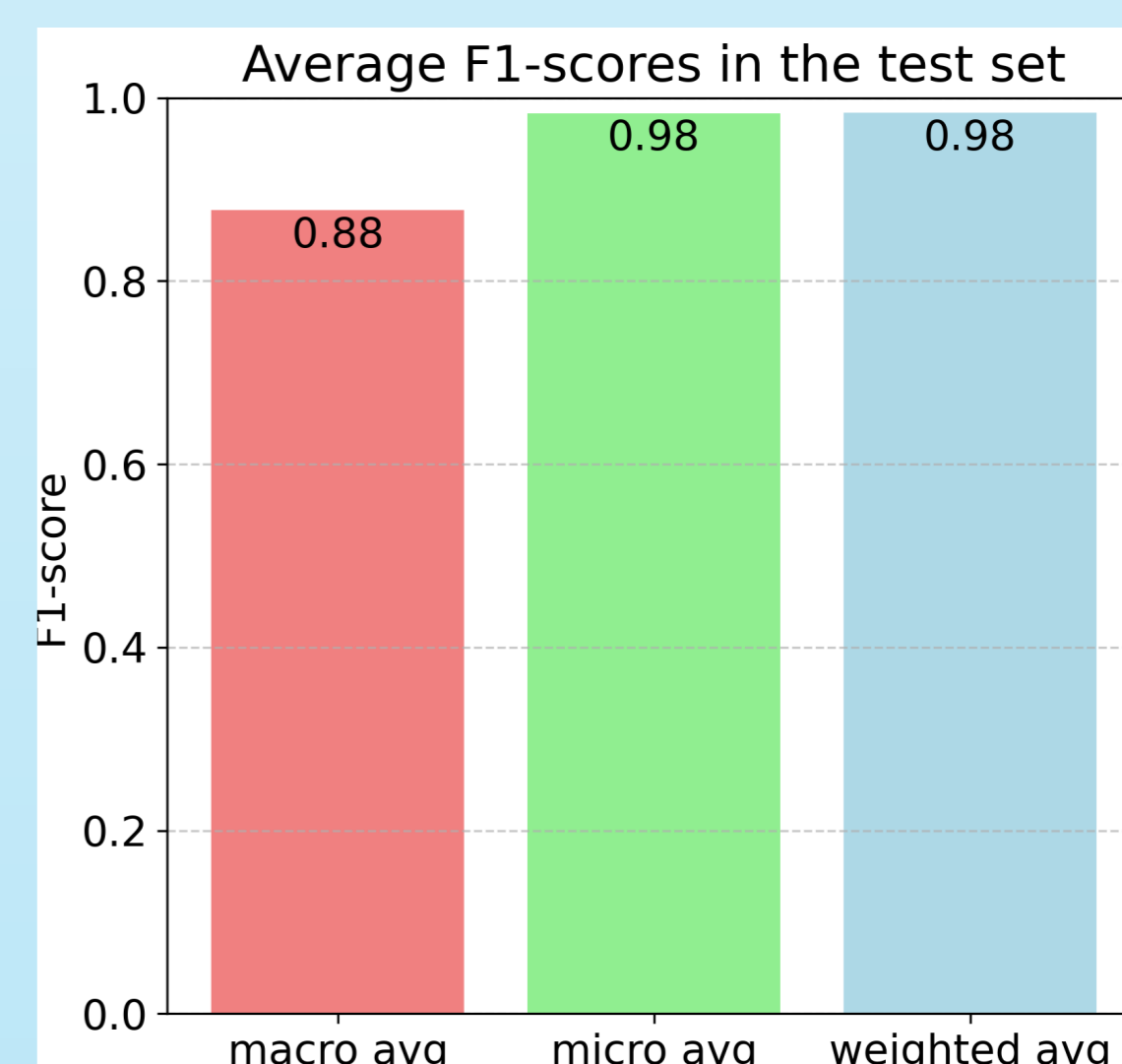
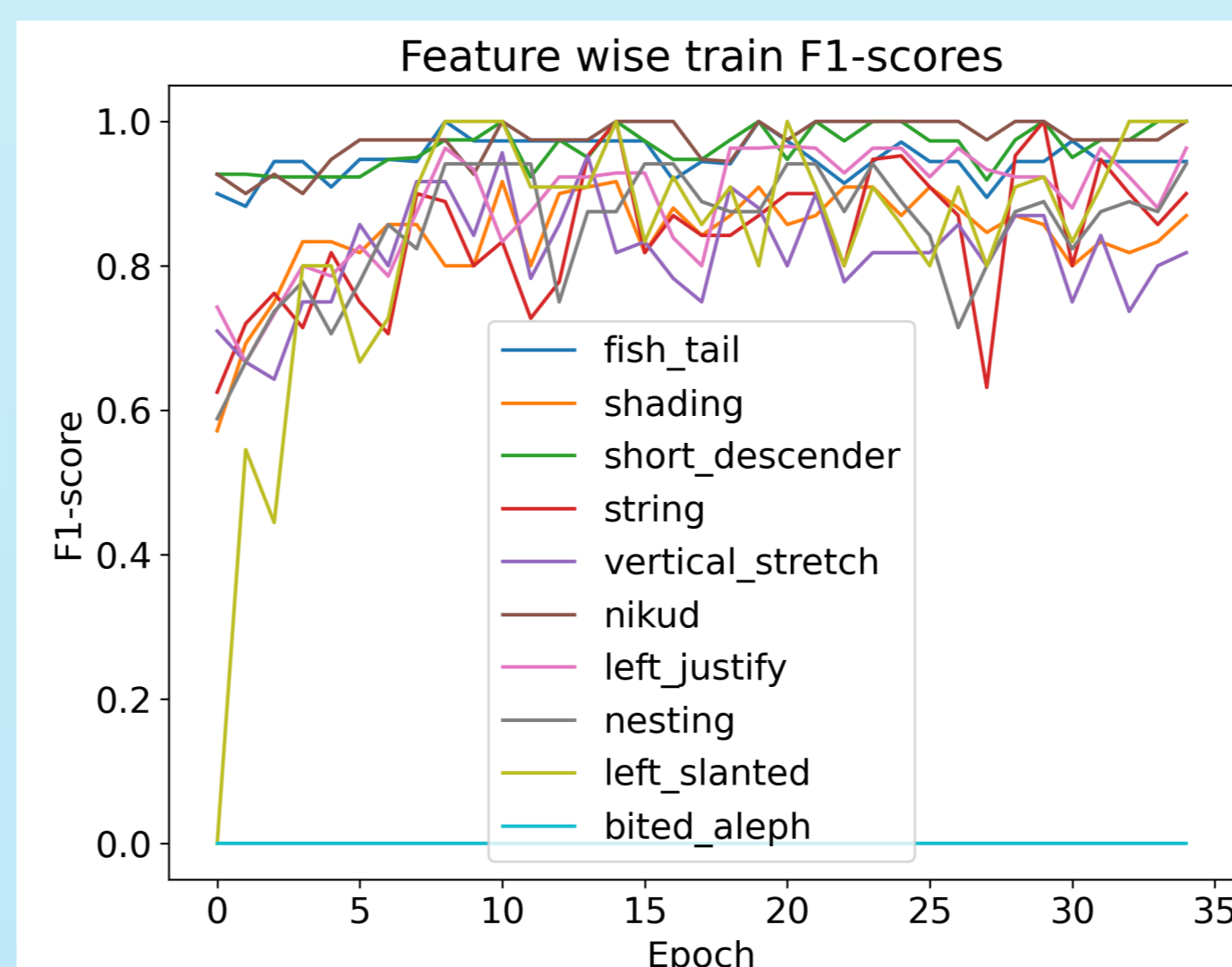
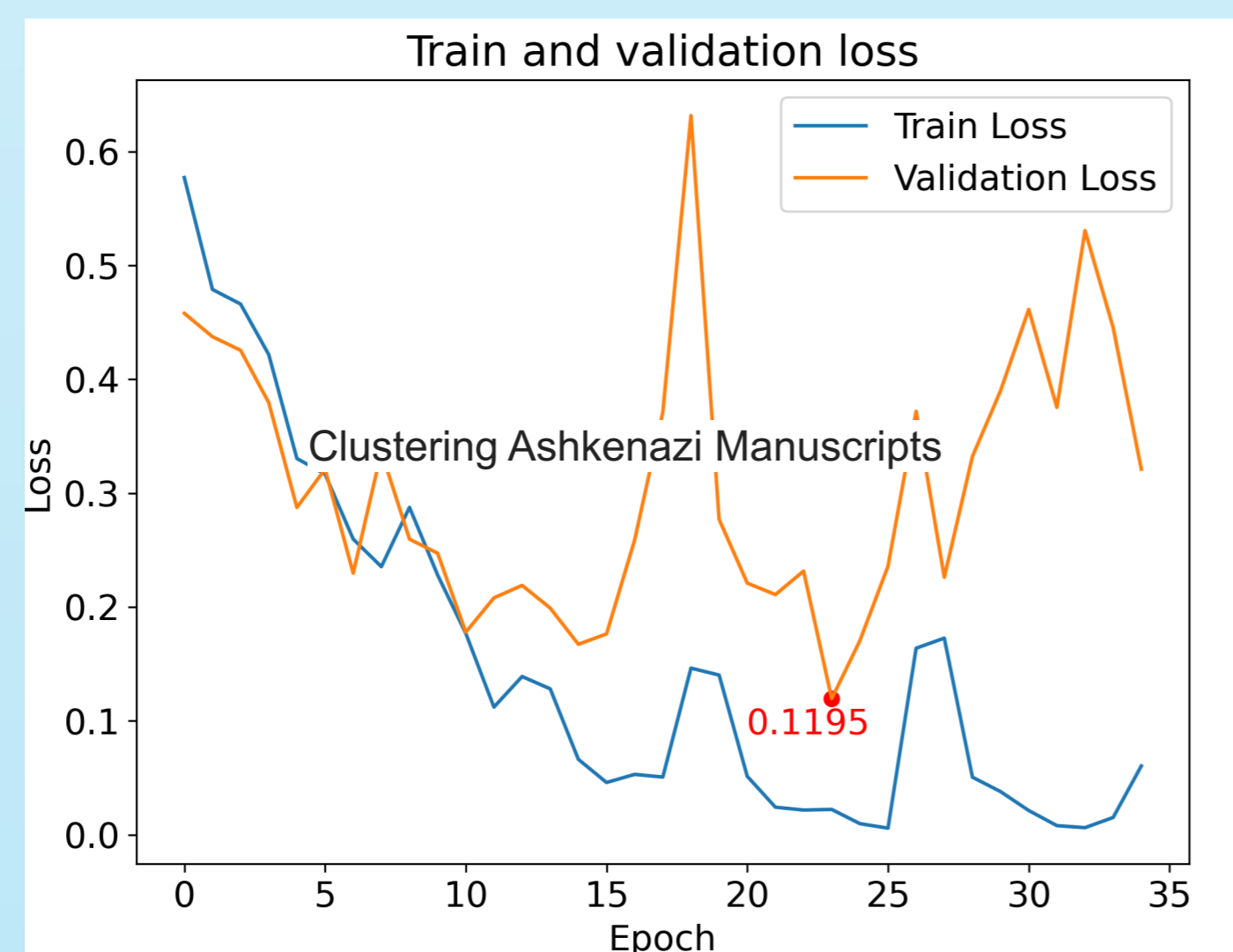
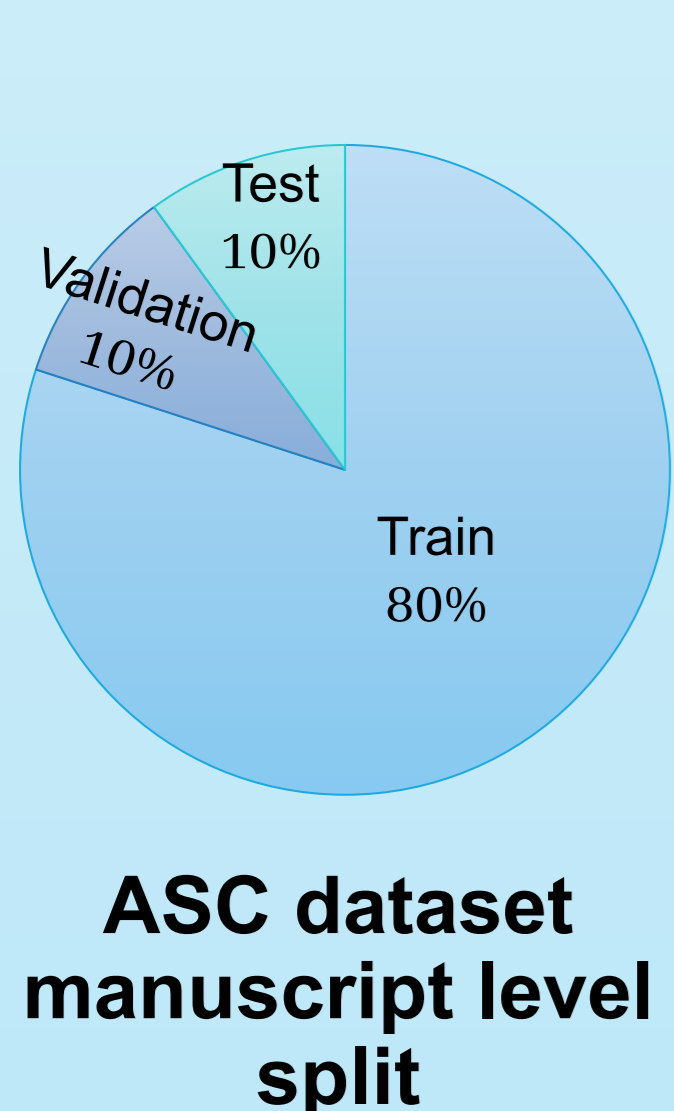
Unknown samples



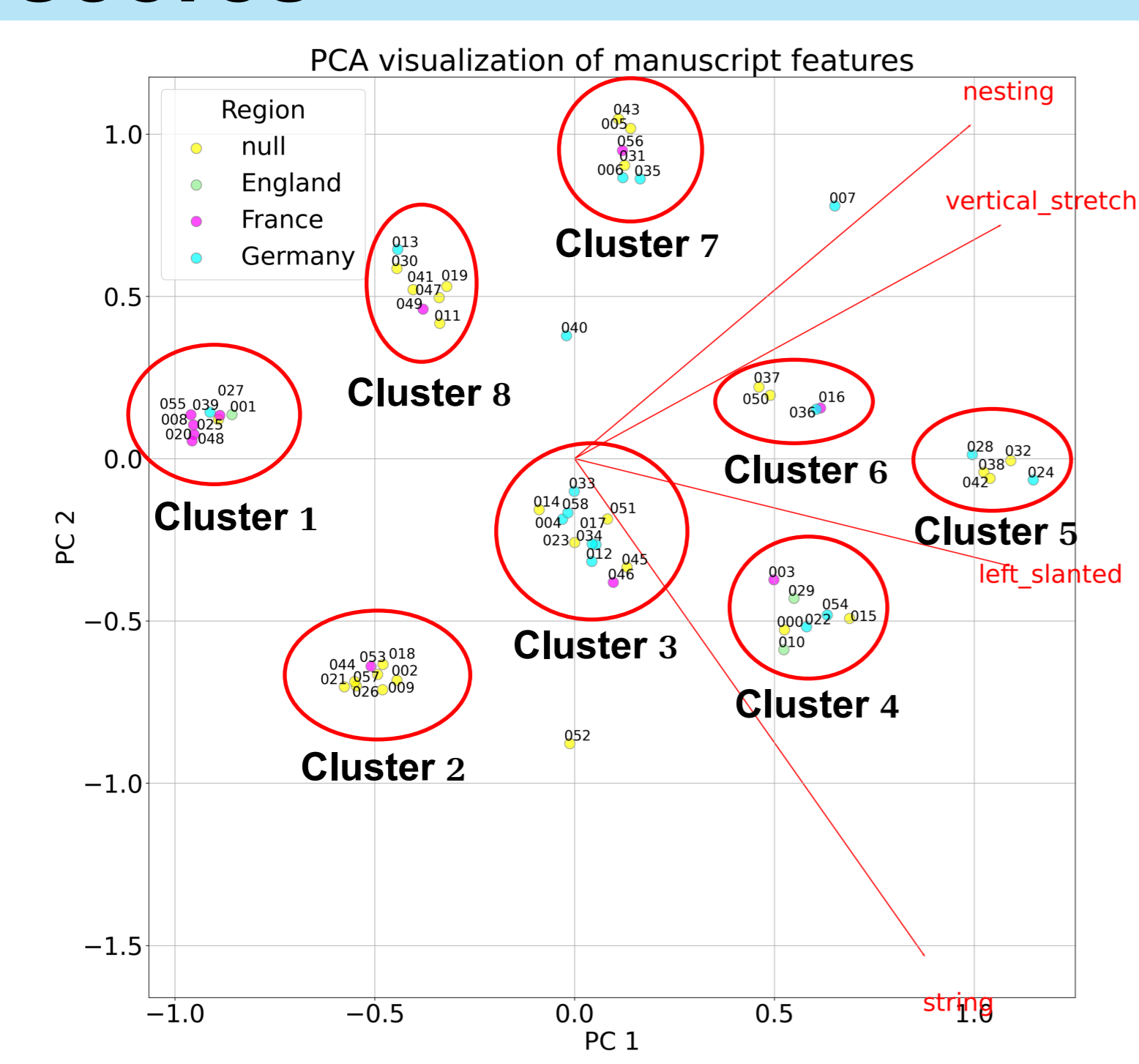
ASC dataset summary

	German	French	Unknown	Total
Manuscripts	17	11	31	59
Pages	62	35	109	206
Text regions	136	61	260	457
Text lines	4413	1799	8080	14292

Training a VGG-19 network for predicting paleographic features



Clusters within Ashkenazi square script using features with the highest χ^2 scores



Cluster 1 [None]

Cluster 2 [string]

Cluster 3 [string, nesting] [string, vertical stretch]

Cluster 4 [string, left slanted, nesting] [string, left slanted, vertical stretch]

Cluster 5 [string, left slanted, nesting, vertical stretch]

Cluster 6 [string, nesting, vertical stretch]

Cluster 7 [vertical stretch, nesting]

Cluster 8 [vertical stretch] [nesting]

