

6. Create a TF-IDF + Logistic Regression spam detector on the SMS spam dataset. Report precision, recall, F1 and list the top positive/negative

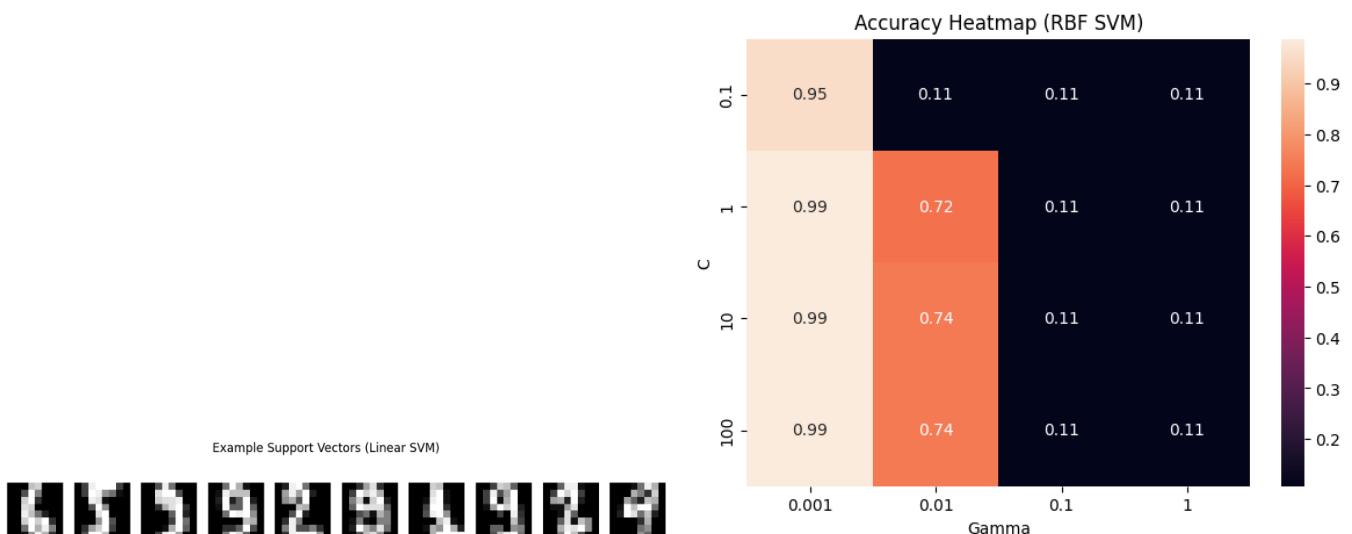
	label	message
0	ham	Go until jurong point, crazy.. Available only ...
1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...

	label	message
0	ham	Go until jurong point, crazy.. Available only ...
1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...

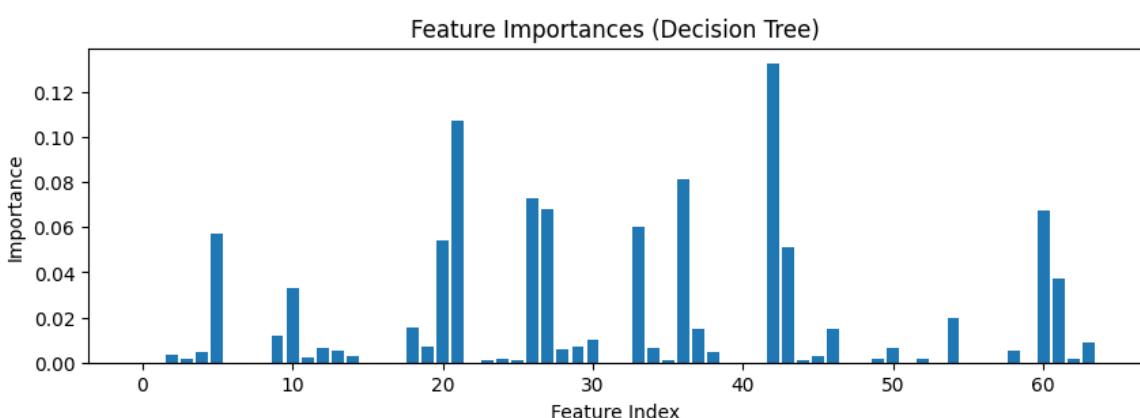
Top spam indicators (positive weights): ['18' 'com' 'won' 'cash' 'win' 'urgent' 'new' 'prize' '50' '150p' 'service' 'text' 'free' 'reply' 'uk' 'stop' 'www' 'claim' 'mobile' 'txt']

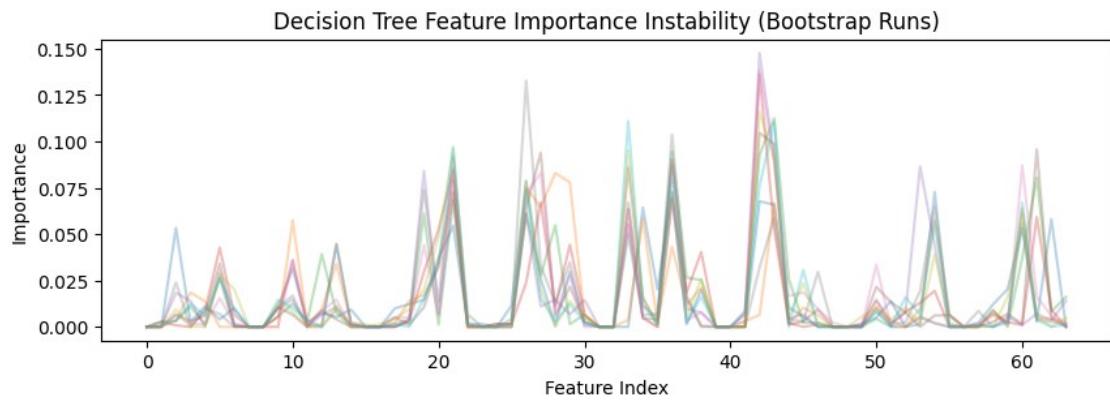
Top ham indicators (negative weights): ['ok' 'll' 'lt' 'gt' 'come' 'da' 'home' 'sorry' 'got' 'later' 'lor' 'going' 'oh' 'way' 'good' 'did' 'doing' 'wat' 'think' 'sir']

7. SVM Kernels and margins Train Linear and RBF SVM on digits;tune C and gamma via grid search, visualize support vectors and plot accuracy heatmap over parameters C and gamma

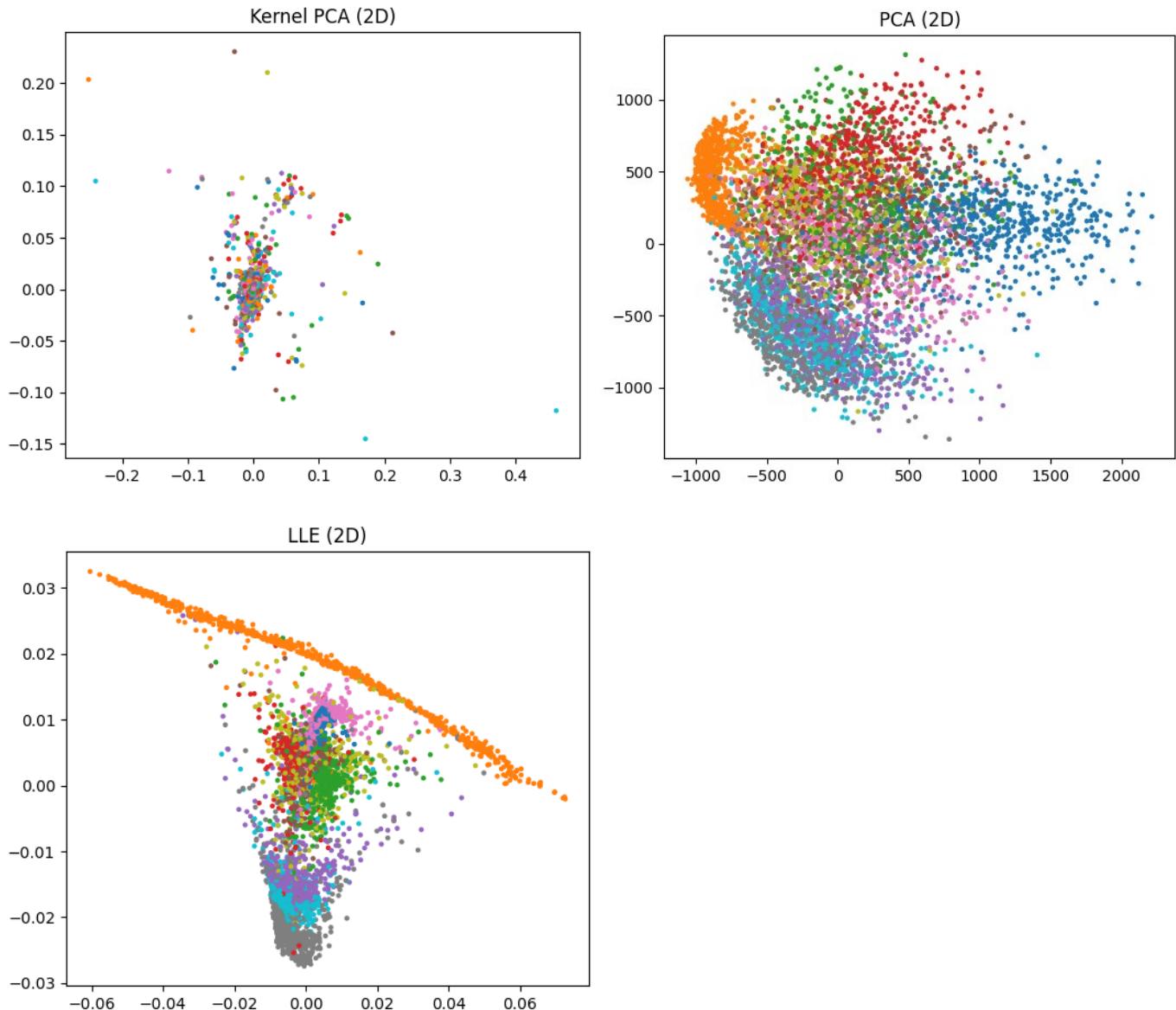


8. Fit and visualize a DecisionTreeClassifier (gini vs entropy) with depth/min samples tuning. Show feature importances and demonstrate instability via bootstraps

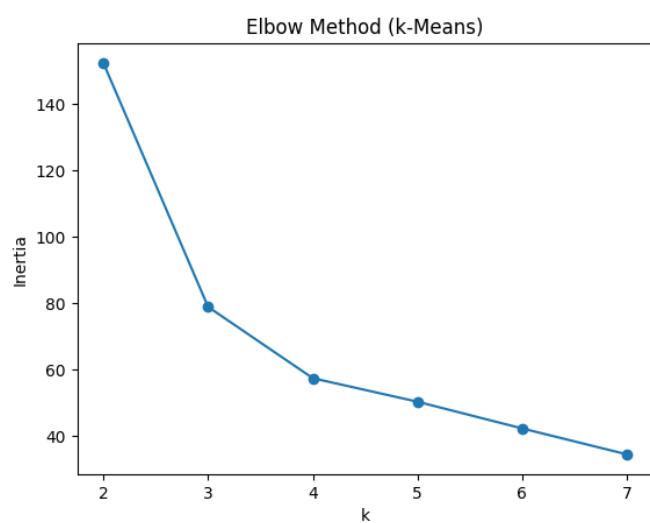
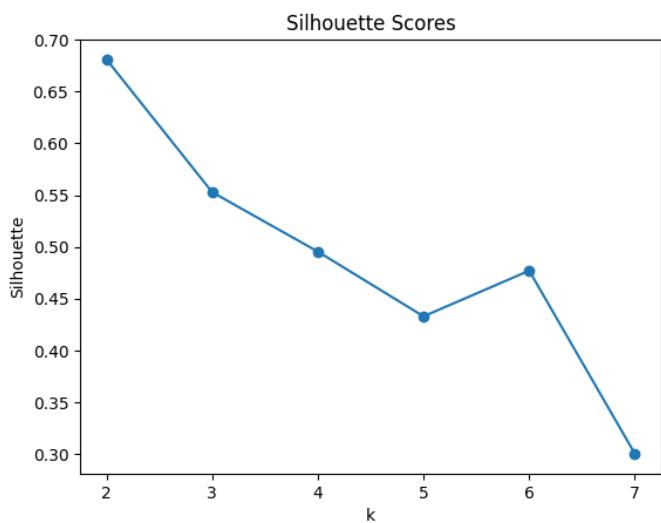




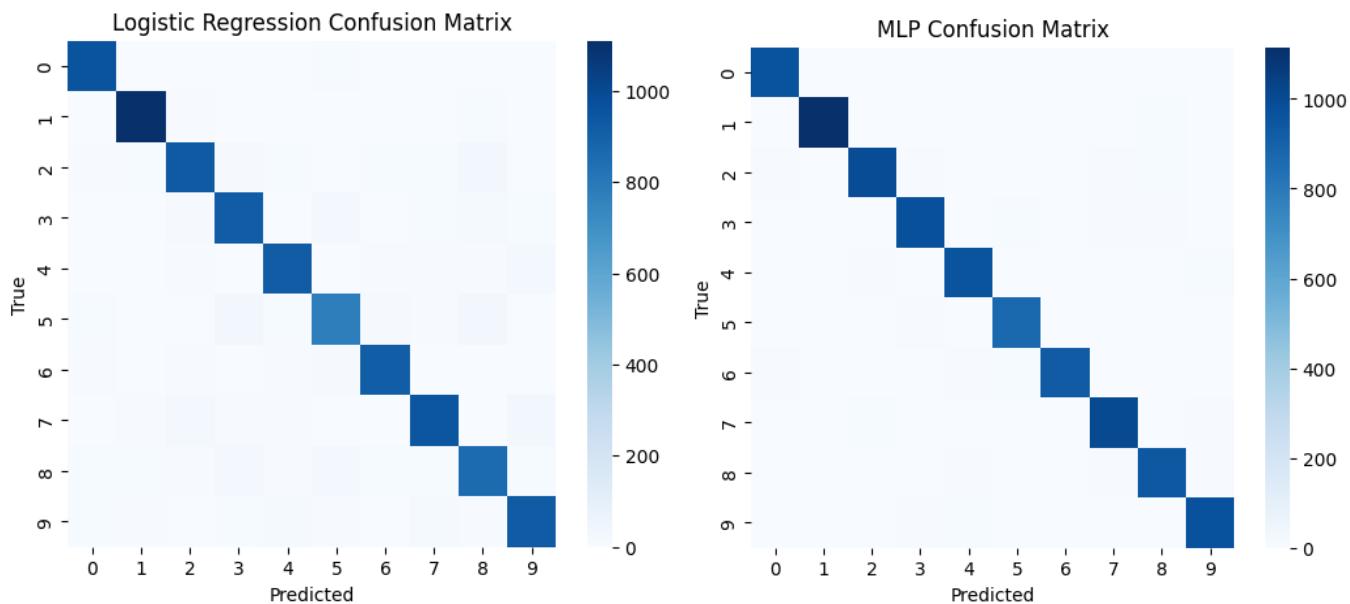
10. Dimensionality Reduction (PCA, KPCA, LLE) Reduce Digits to 2D with PCA, Kernel PCA, and LLE; visualize clusters. Evaluate k-NN accuracy before/after reduction.



11. Unsupervised Clustering (k-Means vs GMM) Cluster Iris/Digits using k-Means and Gaussian Mixtures; choose k via elbow/silhouette. Compare cluster quality with silhouette score and ARI (using labels only for evaluation).



12. MNIST with Keras (MLP)



13. NLP Mini-Suite: Information Retrieval + Information Extraction

```
Top 5 Results for Query: 'AI company in California'
#1 | Doc 3 | Score: 0.472
Text: Google unveiled new AI tools at the California headquarters.
Entities: [('Google', 'ORG'), ('AI', 'GPE'), ('California', 'GPE')]

#2 | Doc 5 | Score: 0.099
Text: Microsoft acquired several startups in 2022.
Entities: [('Microsoft', 'ORG'), ('2022', 'DATE')]

#3 | Doc 7 | Score: 0.089
Text: Amazon is launching new delivery drones in the USA.
Entities: [('Amazon', 'ORG'), ('USA', 'GPE')]

#4 | Doc 4 | Score: 0.083
Text: The Eiffel Tower in Paris attracts millions of tourists.
Entities: [('The Eiffel Tower', 'LOC'), ('Paris', 'GPE'), ('millions', 'QUANTITY')]

#5 | Doc 2 | Score: 0.079
Text: The Indian cricket team won the match held in Mumbai.
Entities: [('Indian', 'NORP'), ('Mumbai', 'GPE')]
```