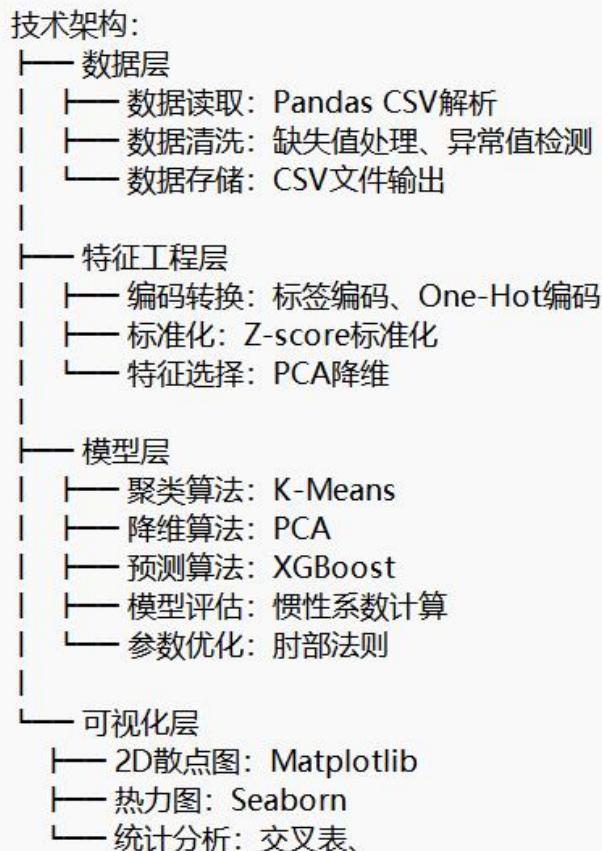


数据集使用的飞桨社区的汽车客户数据集，训练集总计 8068 个样本，测试集总计 2627，我使用了其中的这些数据:Gender : 客户的性别, Married : 客户的婚姻状况, Age : 客户年龄,

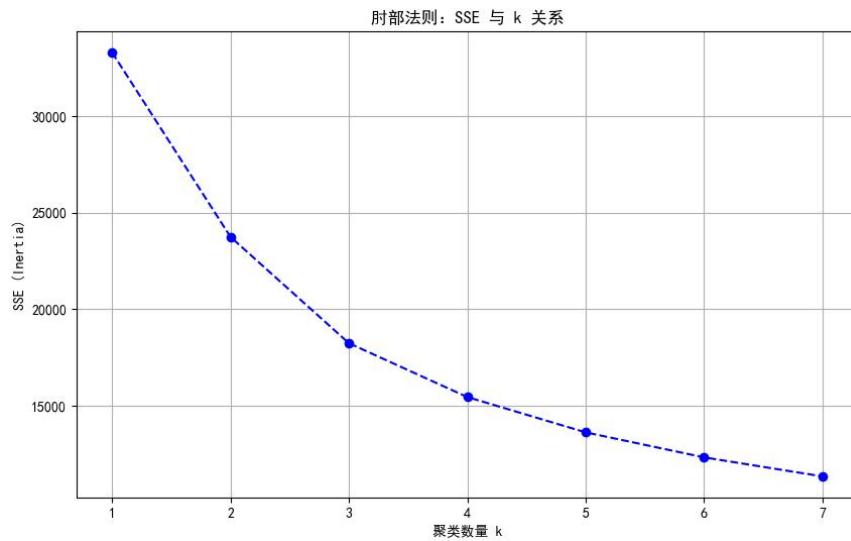
Profession: 客户的职业 WorkExperience : 客户多年的工作经验, SpendingScore : 客户的消费分数

Gender	Married	Age	Profession	WorkExperience	SpendingScore
Male	No	22	Healthcare	1	Low
Female	Yes	38	Engineer		Average
Female	Yes	67	Engineer	1	Low
Male	Yes	67	Lawyer	0	High
Female	Yes	40	Entertainment		High
Male	Yes	56	Artist	0	Average
Male	No	32	Healthcare	1	Low
Female	No	33	Healthcare	1	Low
Female	Yes	61	Engineer	0	Low
Female	Yes	55	Artist	1	Average
Female	No	26	Engineer	1	Low
Male	No	19	Healthcare	4	Low
Female	No	19	Executive	0	Low
Male	Yes	70	Lawyer		Low
Female	Yes	58	Doctor	0	Low
Female	No	41	Healthcare	1	Low
Female	No	32	Homemaker	9	Low
Male	No	31	Healthcare	1	Low

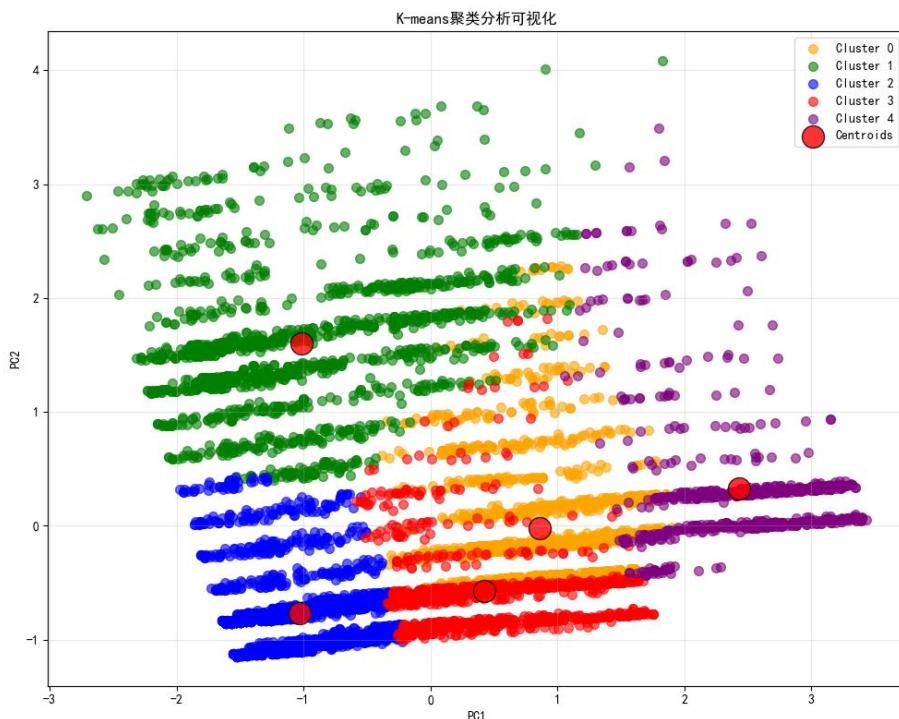
技术架构：



肘部法则：



训练集聚类分析可视化：



PCA 分析成分:

主要由前三个主导，由前三个就可以分析用户画像

聚类用户画像分析:

Cluster	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	
0	0.86	-0	-1	-0	0.1	-0	-0	0.0	-0	PC1较高, PC2低, 年轻高消费 + 非传统路径
1	-1	1.60	0.22	0.00	#	-0	0.00	0.0	-0	PC1低, PC2高, 年长 + 工龄长 + 低消费
2	-1	-0.8	-0	-0	#	0.06	0.03	-0	0.0106272919775	PC1低, PC2低, 年轻低中龄低消费, 学生
3	0.42	-0.6	1.14	0.11	0.1	-0	-0	0.0	-0	PC3高, PC1略高, 高龄 + 低消费 + 工龄短
4	2.43	0.32	-0	0.06	#	0.26	0.09	-0	0.0295245880072	极端高消费 + 非传统路径

