



&



Comparative Study

b/w

Python Code and SKLearn Libraries

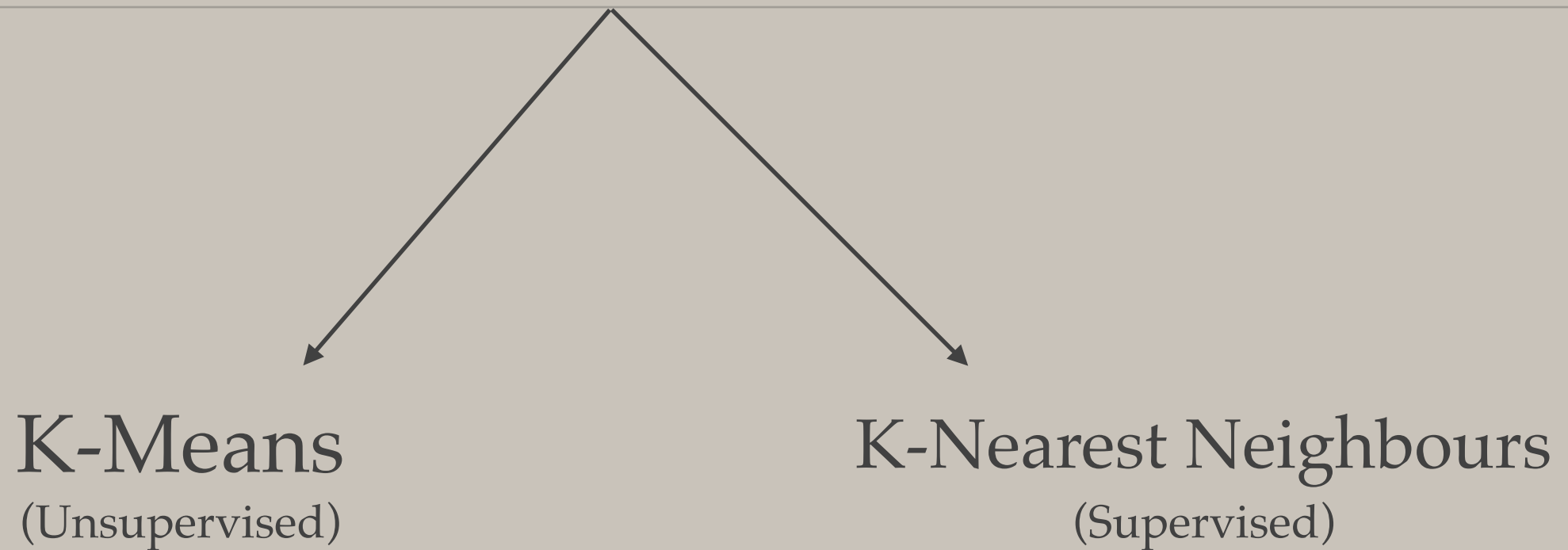
Group Members :-

Siddhanth J Ajri

Venkatesh B.N

Yugesh Kumar

Algorithms Covered

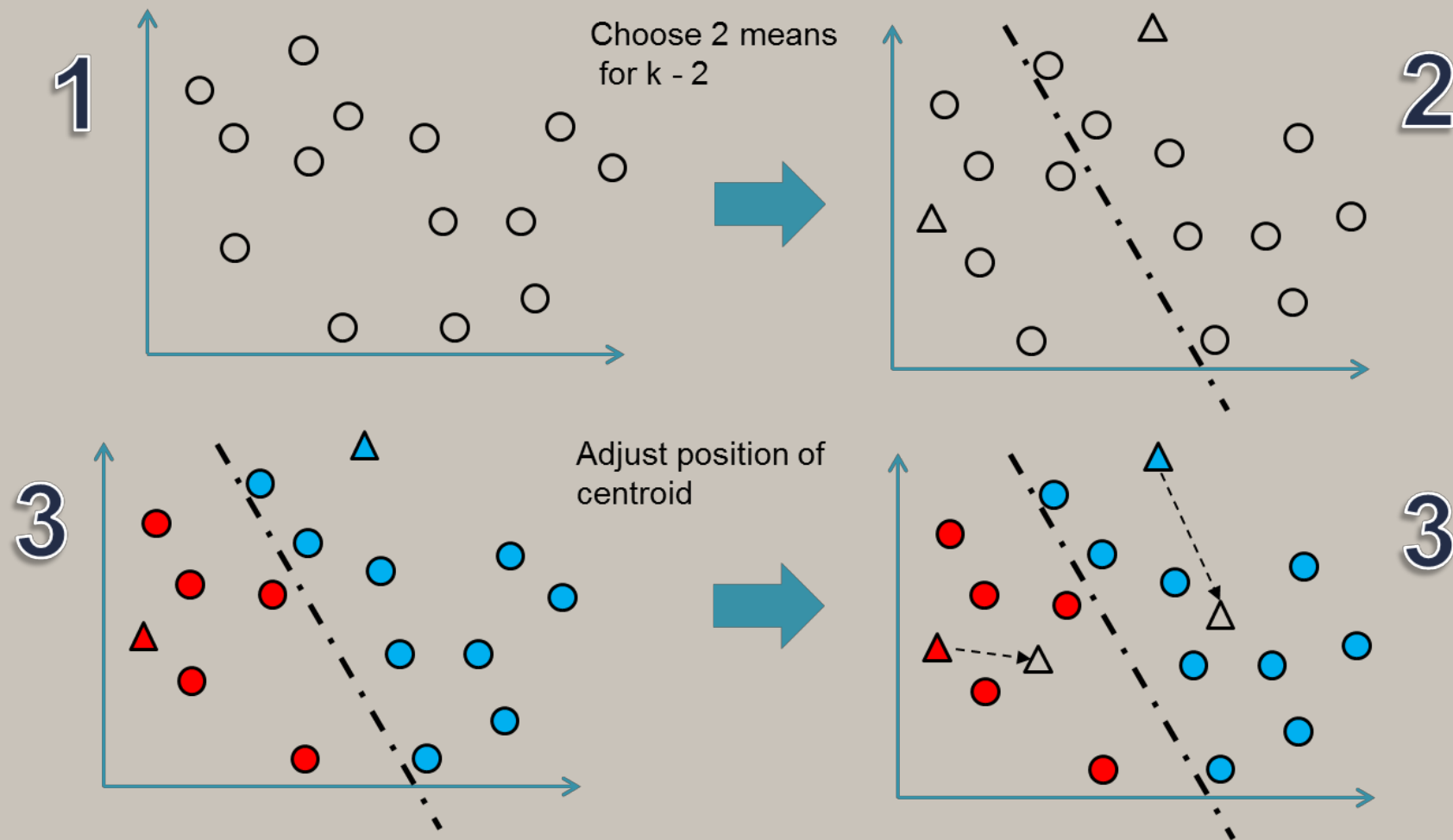


Both the algorithms are classifying the data
based on specific set of attributes...

Sample of Dataset used...

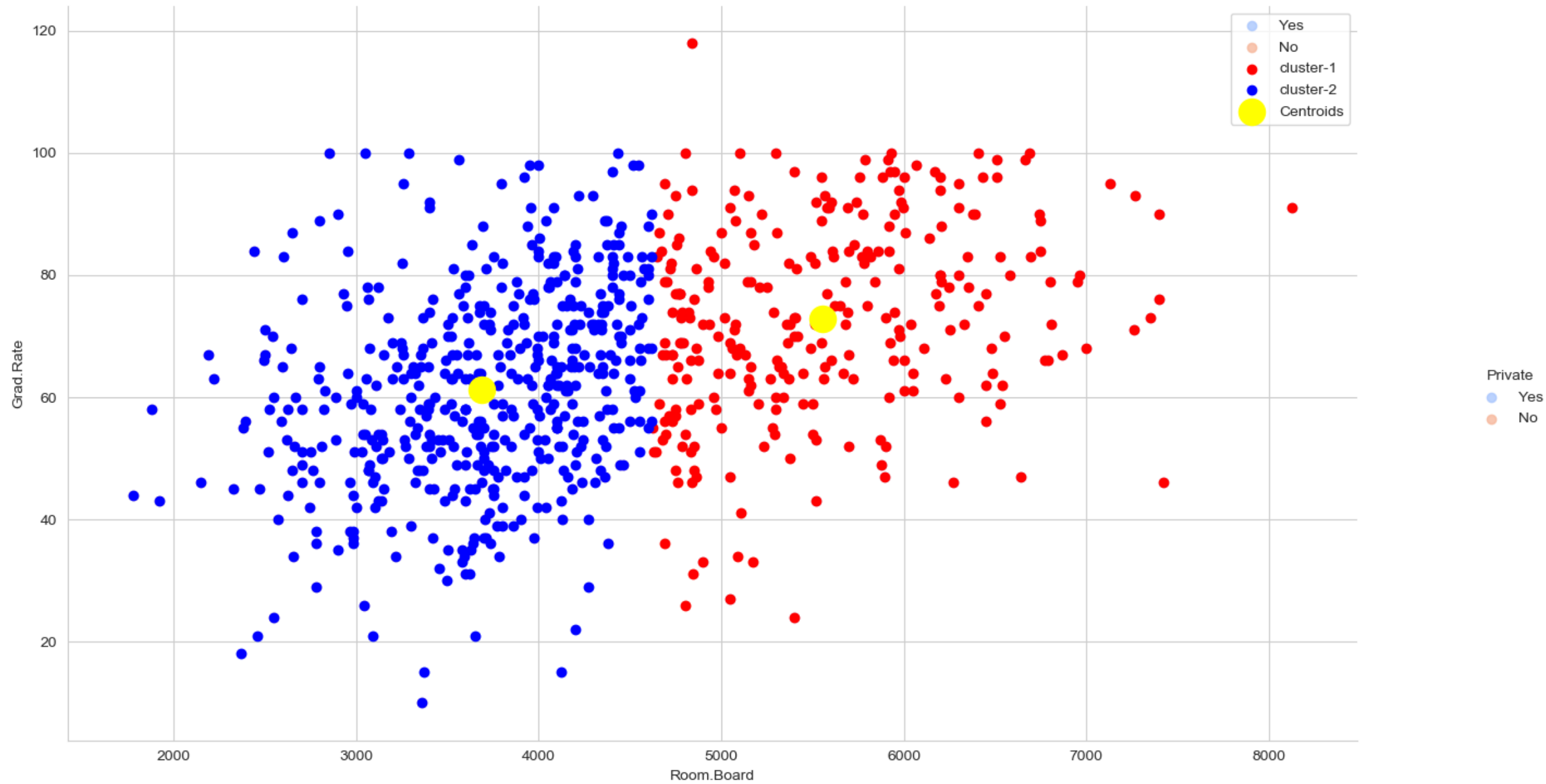
	Private	Apps	Accept	Enroll	Top10perc	Top25perc	F.Undergrad	P.Undergrad	Outstate	Room.Board	Books	Personal	PhD	Terminal	S.F.Ratio	perc.alumni	Expend	Grad.Rate
Abilene Christian University	Yes	1660	1232	721	23	52	2885	537	7440	3300	450	2200	70	78	18.1	12	7041	60
Adelphi University	Yes	2186	1924	512	16	29	2683	1227	12280	6450	750	1500	29	30	12.2	16	10527	56
Adrian College	Yes	1428	1097	336	22	50	1036	99	11250	3750	400	1165	53	66	12.9	30	8735	54
Agnes Scott College	Yes	417	349	137	60	89	510	63	12960	5450	450	875	92	97	7.7	37	19016	59
Alaska Pacific University	Yes	193	146	55	16	44	249	869	7560	4120	800	1500	76	72	11.9	2	10922	15
Albertson College	Yes	587	479	158	38	62	678	41	13500	3335	500	675	67	73	9.4	11	9727	55
Albertus Magnus College	Yes	353	340	103	17	45	416	230	13290	5720	500	1500	90	93	11.5	26	8861	63
Albion College	Yes	1899	1720	489	37	68	1594	32	13868	4826	450	850	89	100	13.7	37	11487	73
Albright College	Yes	1038	839	227	30	63	973	306	15595	4400	300	500	79	84	11.3	23	11644	80
Alderson-Broaddus College	Yes	582	498	172	21	44	799	78	10468	3380	660	1800	40	41	11.5	15	8991	52
Alfred University	Yes	1732	1425	472	37	75	1830	110	16548	5406	500	600	82	88	11.3	31	10932	73
Allegheny College	Yes	2652	1900	484	44	77	1707	44	17080	4440	400	600	73	91	9.9	41	11711	76
Allentown Coll. of St. Francis de Sales	Yes	1179	780	290	38	64	1130	638	9690	4785	600	1000	60	84	13.3	21	7940	74
Alma College	Yes	1267	1080	385	44	73	1306	28	12572	4552	400	400	79	87	15.3	32	9305	68
Alverno College	Yes	494	313	157	23	46	1317	1235	8352	3640	650	2449	36	69	11.1	26	8127	55
American International College	Yes	1420	1093	220	9	22	1018	287	8700	4780	450	1400	78	84	14.7	19	7355	69
Amherst College	Yes	4302	992	418	83	96	1593	5	19760	5300	660	1598	93	98	8.4	63	21424	100
Anderson University	Yes	1216	908	423	19	40	1819	281	10100	3520	550	1100	48	61	12.1	14	7994	59
Andrews University	Yes	1130	704	322	14	23	1586	326	9996	3090	900	1320	62	66	11.5	18	10908	46
Angelo State University	No	3540	2001	1016	24	54	4190	1512	5130	3592	500	2000	60	62	23.1	5	4010	34
Antioch University	Yes	713	661	252	25	44	712	23	15476	3336	400	1100	69	82	11.3	35	42926	48
Appalachian State University	No	7313	4664	1910	20	63	9940	1035	6806	2540	96	2000	83	96	18.3	14	5854	70
Aquinas College	Yes	619	516	219	20	51	1251	767	11208	4124	350	1615	55	65	12.7	25	6584	65
Arizona State University Main campus	No	12809	10308	3761	24	49	22593	7585	7434	4850	700	2100	88	93	18.9	5	4602	48
Arkansas College (Lyon College)	Yes	708	334	166	46	74	530	182	8644	3922	500	800	79	88	12.6	24	14579	54
Arkansas Tech University	No	1734	1729	951	12	52	3602	939	3460	2650	450	1000	57	60	19.6	5	4739	48
Assumption College	Yes	2135	1700	491	23	59	1708	689	12000	5920	500	500	93	93	13.8	30	7100	88
Auburn University-Main Campus	No	7548	6791	3070	25	57	16262	1716	6300	3933	600	1908	85	91	16.7	18	6642	69
Augsburg College	Yes	662	513	257	12	30	2074	726	11902	4372	540	950	65	65	12.8	31	7836	58
Augustana College IL	Yes	1879	1658	497	36	69	1950	38	13353	4173	540	821	78	83	12.7	40	9220	71
Augustana College	Yes	761	725	306	21	58	1337	300	10990	3244	600	1021	66	70	10.4	30	6871	69

How K-Means Clustering Work?

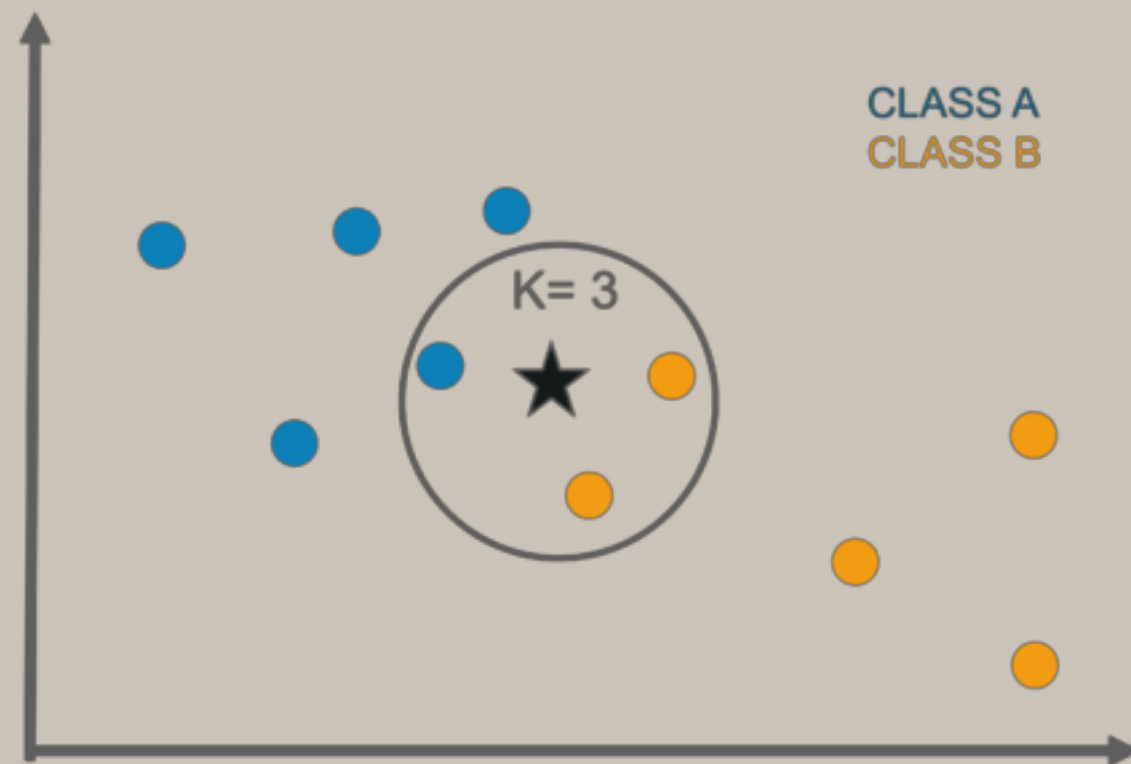


1. First we initialise k points, called means, randomly.
2. We categorise each item to its closest mean and we update the mean's coordinates, which are the averages of the items categorised in the mean so far.
3. We repeat the process for a given number of iterations and at the end, we have our clusters.

K-Means Output



How K-Nearest Neighbour Classification Work?



1. We initialise the value of k , where k represents the number of nearest neighbours.
2. We now make a circle with our object as the centre enclosing k nearest points.
3. We check the class to which the nearest neighbours belong.
4. With confidence, we assign our test object a class, to which most of the other neighbouring objects belong (majority voting scheme).

KNN Error Rate Graph

