

CMP-6002A - Machine Learning

Dr Gavin Cawley

Seminar 2 : k-Nearest Neighbour Classifiers

Question

The table below gives the height and weight of ten sumo wrestlers, classified as sanyaku (title-holders) or maegashira (lesser wrestlers).

Table 1. Training data

Name	Height (cm)	Weight (Kg)	Rank
Hakuho	192	155	sanyaku
Asashoryu	185	147	sanyaku
Chiyotaikai	180	152	sanyaku
Kaio	185	171	sanyaku
Kaiho	177	122	maegashira
Yoshikaze	177	131	maegashira
Iwakiyama	184	174	maegashira
Wakakirin	186	135	maegashira
Kakizoe	178	140	maegashira
Kakuryu	187	138	maegashira

Question (ctd.)

Use the 3-nearest neighbour classifier to predict whether the wrestlers listed in the table below are sanyaku or maegashira

Table 2. Test data

Name	Height (cm)	Weight (Kg)	Rank
Takekaze	171	143	???
Kotomitsuki	181	149	???
Takamisakari	188	140	???
Kotooshu	202	152	???

Takekaze, Height = 171 cm, Weight = 143 kg

using the squared distance metric (the square of the Euclidean distance)

Name	Height	Weight	M/S	Dist	Rank
Hakuho	192	155	sanyaku	585	8
Asashoryu	185	147	sanyaku	212	4
Chiyotaikai	180	152	sanyaku	162	2
Kaio	185	171	sanyaku	980	9
Kaiho	177	122	maegashira	477	7
Yoshikaze	177	131	maegashira	180	3
Iwakiyama	184	174	maegashira	1130	10
Wakakirin	186	135	maegashira	289	5
Kakizoe	178	140	maegashira	58	1
Kakuryu	187	138	maegashira	281	6

The three nearest neighbours are Kakizoe (M), Chiyotaikai (S) and Yoshikaze (M), so Takekaze is classified as Maegashira

Kotomitsuki, Height = 181 cm, Weight = 149 kg

using the squared distance metric (the square of the Euclidean distance)

Name	Height	Weight	M/S	Dist	Rank
Hakuho	192	155	sanyaku	157	4.5
Asashoryu	185	147	sanyaku	20	2
Chiyotaikai	180	152	sanyaku	10	1
Kaio	185	171	sanyaku	500	8
Kaiho	177	122	maegashira	745	10
Yoshikaze	177	131	maegashira	340	7
Iwakiyama	184	174	maegashira	634	9
Wakakirin	186	135	maegashira	221	6
Kakizoe	178	140	maegashira	90	3
Kakuryu	187	138	maegashira	157	4.5

The three nearest neighbours are Chiyotaikai (S) Asashoryu (S) and Kakizoe (M), so Kotomitsuki is classified as Sanyaku

Takamisakari Height = 188 cm, Weight = 140 kg

using the squared distance metric (the square of the Euclidean distance)

Name	Height	Weight	M/S	Dist	Rank
Hakuho	192	155	sanyaku	241	7
Asashoryu	185	147	sanyaku	58	3
Chiyotaikai	180	152	sanyaku	208	6
Kaio	185	171	sanyaku	970	9
Kaiho	177	122	maegashira	445	8
Yoshikaze	177	131	maegashira	202	5
Iwakiyama	184	174	maegashira	1172	10
Wakakirin	186	135	maegashira	29	2
Kakizoe	178	140	maegashira	100	4
Kakuryu	187	138	maegashira	5	1

The three nearest neighbours are Kakuryu (M), Wakakirin (M) and Asashoryu (S), so Takamisakari is classified as Maegashira

Kotooshu, Height = 202 cm, Weight = 152 kg

using the squared distance metric (the square of the Euclidean distance)

Name	Height	Weight	M/S	Dist	Rank
Hakuho	192	155	sanyaku	109	1
Asashoryu	185	147	sanyaku	314	2
Chiyotaikai	180	152	sanyaku	484	4
Kaio	185	171	sanyaku	650	6
Kaiho	177	122	maegashira	1525	10
Yoshikaze	177	131	maegashira	1066	9
Iwakiyama	184	174	maegashira	808	8
Wakakirin	186	135	maegashira	545	5
Kakizoe	178	140	maegashira	720	7
Kakuryu	187	138	maegashira	421	3

The three nearest neighbours are Hakuho (S), Asashoryu (S) and Kakuryu (M), so Kotooshu is classified as Sanyaku

Conclusions

- ▶ This question was set by an sumo fan!
- ▶ The k-NN classifier is pretty straightforward
 - ▶ Easy marks available **IF** it turns up on the exam!
- ▶ Homework - see what difference the measuring height in meters instead of centimeters makes