

Verhaltensklassifikation

von Tim Korinth

Aufgabe:

Das Thema dieser Aufgabe ist die Klassifikation von Verhaltensmuster.

Anhand der individuellen Bewegungsdaten von Fischen soll bestimmt werden, ob das Individuum alleine oder in einem Schwarm unterwegs war. Mit Hilfe der Daten aus train_alone.txt und train_group.txt, welche Bewegungsdaten der Fische beinhalten, soll ein Algorithmus implementiert werden, der mit einer Zuverlässigkeit von jeweils mindestens 0.9 entscheiden kann, zu welcher Kategorie ein Fisch gehört. Verwendet werden sollten hierzu die Bayes'sche Klassifikation eine Markovkette erster Ordnung.

Lösung:

Zunächst einmal werden wie die Bewegungsdaten der Fische aus beiden Trainings-Textdateien ausgelesen und durch deren Umwandlung in Differenzen für beide Textdateien jeweils ein 9x9-Array erstellt, das die Wahrscheinlichkeiten der Folgezustände für jeden Zustand beinhaltet. Dieses wird anschließend logarithmiert. Um Den Wert 0 hierfür auszuschließen, werden die Arrays jeweils vorher mit 1 initialisiert.

Die Daten aus den Evaluationsdateien werden anschließend pro Fisch bestimmt, indem für jeden Übergang eines Zustand in den anderen dieWahrscheinlichkeiten in einer Gruppe und alleine addiert werden. Der höhere Wert gibt an, welche soziale Form wahrscheinlicher ist.

Beim Testen des Algorithmus stellte sich heraus, dass die höchste Genauigkeit bei der Logarithmierung mit dem natürlichen Logarithmus aus der Summe des Wertes und 1 (in Java: `Math.log1p(<Wert>)`) erreicht wurde, während der natürliche Logarithmus und der Logarithmus zur Basis 10 unsicherere Ergebnisse lieferten.

Ergebnisse:

Zunächst wurde in Schritten von 0,1 grob bestimmt, wo beide Ausgaben einen Wert von 0.75 oder größer erreicht haben. Anschließend wurden diese Bereiche durch kleinere Schritte genauer untersucht um Bereiche mit den höchsten Werten zu bestimmen, welche genauer untersucht werden. Durch genauere Eingrenzung stellte sich der Bereich von **3.5880** bis **3.5969** als der geeigneteste heraus, da hier die Effizienz am höchsten war:

Genauigkeit der Bestimmung für einzelne Fische:

0.89583

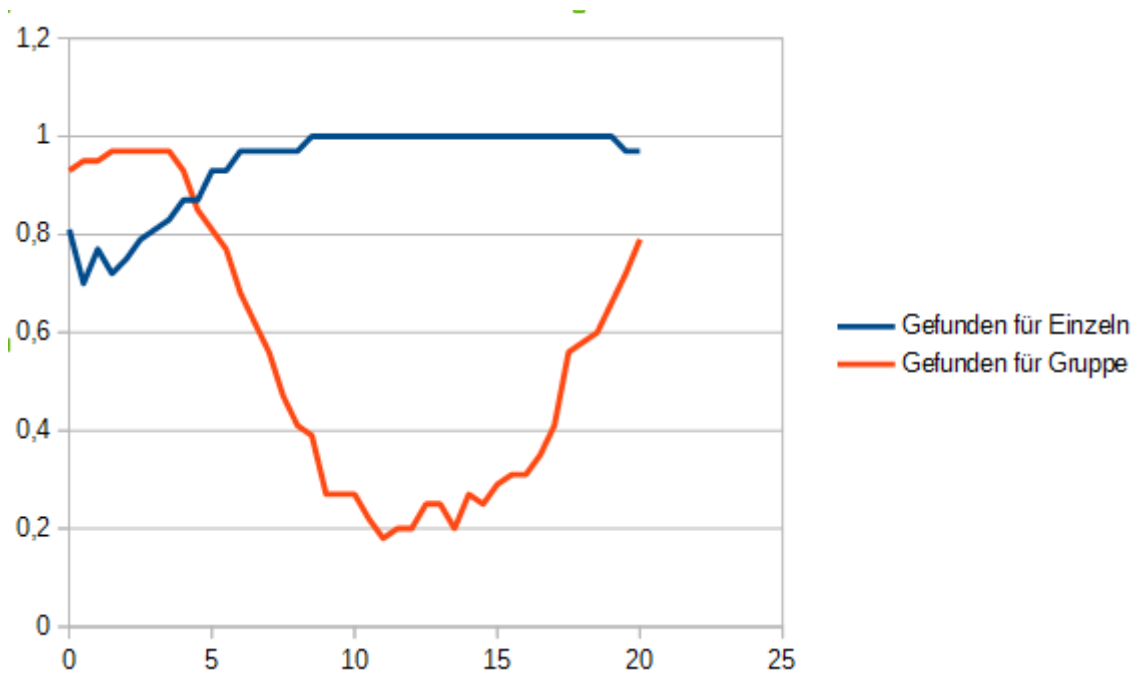
Genauigkeit der Bestimmung für Fische in Gruppen:

0.97916

Um die bestmöglichen Werte zu bestimmen war es notwendig, den Schwellwert k in Schritten von 0,00001 zu prüfen. Das nachfolgende Diagramm wurde allerdings aus praktischen Gründen mit Daten aus 0,5er Schritten erstellt. Daher lässt sich das Maximum bei 3,59 nicht deutlich erkennen.

Diagramm:

Diagramm für beide Genauigkeiten abhängig vom Wert k (0,5er Schritte). Es lassen sich Maxima bei ~ 4 und ~ 21 erkennen.



Wert für k	Gefunden für Einzel	Gefunden für Gruppe
0	0,81	0,93
0,5	0,7	0,95
1	0,77	0,95
1,5	0,72	0,97
2	0,75	0,97
2,5	0,79	0,97
3	0,81	0,97
3,5	0,83	0,97
4	0,87	0,93
4,5	0,87	0,85
5	0,93	0,81
5,5	0,93	0,77
6	0,97	0,68
6,5	0,97	0,62
7	0,97	0,56
7,5	0,97	0,47
8	0,97	0,41
8,5	1	0,39
9	1	0,27
9,5	1	0,27
10	1	0,27
10,5	1	0,22
11	1	0,18
11,5	1	0,2
12	1	0,2
12,5	1	0,25
13	1	0,25
13,5	1	0,2
14	1	0,27
14,5	1	0,25
15	1	0,29
15,5	1	0,31
16	1	0,31
16,5	1	0,35
17	1	0,41
17,5	1	0,56
18	1	0,58
18,5	1	0,6
19	1	0,66
19,5	0,97	0,72
20	0,97	0,79

Anhang:

Wahrscheinlichkeit für Übergang von Zustand zu Zustand:

Wobei

0 = AA, 1 = AB, 2 = AC, 3 = BA, 4 = BB, 5 = BC, 6 = CA, 7 = CB, 8 = CC

Wahrscheinlichkeit	EINZELN	GRUPPE
Zustand 0zu0:	0.1306955606272694	0.16351127274951957
Zustand 0zu1:	0.05753455037510502	0.06491922374998212
Zustand 0zu2:	0.13898007731950363	0.1397594897145738
Zustand 0zu3:	0.06713610484852409	0.07129555289200443
Zustand 0zu4:	0.03864281809620096	0.03704434864603015
Zustand 0zu5:	0.07937952113275902	0.07323725012507341
Zustand 0zu6:	0.1374788975394124	0.143480606425101
Zustand 0zu7:	0.06286727806026877	0.06078720536897815
Zustand 0zu8:	0.2218129447063367	0.18291369824846843
Zustand 1zu0:	0.12412827447353905	0.14820340507881513
Zustand 1zu1:	0.08217838961502935	0.0844023876033468
Zustand 1zu2:	0.13459435555634156	0.13552626119001018
Zustand 1zu3:	0.07167409134815333	0.07469296134810464
Zustand 1zu4:	0.06132505377809649	0.054081593526073425
Zustand 1zu5:	0.06239143617720529	0.06859228152183557
Zustand 1zu6:	0.1483808860790319	0.1458460253921032
Zustand 1zu7:	0.09952959534703311	0.08615765385139755
Zustand 1zu8:	0.158355424438949	0.1449015152864677
Zustand 2zu0:	0.13410025321213676	0.15200218603279295
Zustand 2zu1:	0.06210298678688951	0.06296155743736442
Zustand 2zu2:	0.15702417504687535	0.1612056862899626
Zustand 2zu3:	0.07728132838595532	0.07382012781795015
Zustand 2zu4:	0.03730609068431526	0.03784052101439892
Zustand 2zu5:	0.064986525923782	0.06776031573631768
Zustand 2zu6:	0.20432644028578523	0.18142896952748358
Zustand 2zu7:	0.06081874003503916	0.0599302679017647
Zustand 2zu8:	0.13757748089737218	0.13971439119819554
Zustand 3zu0:	0.1262635124747649	0.14292462341090115
Zustand 3zu1:	0.05757920071760699	0.06355406751960059
Zustand 3zu2:	0.15166478123376648	0.14809105318390337
Zustand 3zu3:	0.0983843671618998	0.0920454174712205
Zustand 3zu4:	0.0624055661874296	0.05478065624796619
Zustand 3zu5:	0.12645987968207237	0.10888223174467405
Zustand 3zu6:	0.1290091505988969	0.1384320286235249
Zustand 3zu7:	0.04976583608172943	0.055324459622072185
Zustand 3zu8:	0.14069152645325356	0.1380317083619738
Zustand 4zu0:	0.08051659320520198	0.09970148890850773
Zustand 4zu1:	0.04665004879353148	0.05875300713167504
Zustand 4zu2:	0.07072259530906533	0.08668323916880193
Zustand 4zu3:	0.06393314427199559	0.07444120829626492
Zustand 4zu4:	0.39469724833359465	0.30319971875314267
Zustand 4zu5:	0.053348838758644004	0.0676504718923617

Zustand 4zu6:	0.07092761837487832	- = -	0.08948716650153615
Zustand 4zu7:	0.04265130622436351	- = -	0.05558561482585481
Zustand 4zu8:	0.07072259530906533	- = -	0.08788588513456232
Zustand 5zu0:	0.14675610433662523	- = -	0.1517004488923668
Zustand 5zu1:	0.059141999987529864	- = -	0.06158985260249782
Zustand 5zu2:	0.13843658879706688	- = -	0.1487317207049359
Zustand 5zu3:	0.12857188502132144	- = -	0.11176395269303326
Zustand 5zu4:	0.06117475054632284	- = -	0.056543584534873514
Zustand 5zu5:	0.08481018745899015	- = -	0.0899316073376356
Zustand 5zu6:	0.147584261733854	- = -	0.13458662914741978
Zustand 5zu7:	0.05188079268092055	- = -	0.05214945882341081
Zustand 5zu8:	0.12349680540346479	- = -	0.13469060627254595
Zustand 6zu0:	0.13800981196190254	- = -	0.14733821163530855
Zustand 6zu1:	0.06484015918156522	- = -	0.0626204605095401
Zustand 6zu2:	0.2100132226017252	- = -	0.17559522615261183
Zustand 6zu3:	0.07592273270016968	- = -	0.07339566831774888
Zustand 6zu4:	0.03652740805700872	- = -	0.03603019135672172
Zustand 6zu5:	0.07614095419585289	- = -	0.07258750039035364
Zustand 6zu6:	0.1549912561758627	- = -	0.16817693932691963
Zustand 6zu7:	0.05184647311957695	- = -	0.05967639557639066
Zustand 6zu8:	0.12687223163716793	- = -	0.14141865988600855
Zustand 7zu0:	0.14173259913341724	- = -	0.14685371344219295
Zustand 7zu1:	0.09790582509416795	- = -	0.09192241316561836
Zustand 7zu2:	0.1585801697060286	- = -	0.1447958349024521
Zustand 7zu3:	0.06758313578619525	- = -	0.07093542794628908
Zustand 7zu4:	0.055015383051924244	- = -	0.05387923367690328
Zustand 7zu5:	0.0690343042312552	- = -	0.06788114263426097
Zustand 7zu6:	0.13578181067744008	- = -	0.14440951077170566
Zustand 7zu7:	0.07135180237161773	- = -	0.07770481208897295
Zustand 7zu8:	0.1446947663727415	- = -	0.14363641448271725
Zustand 8zu0:	0.19716434667168412	- = -	0.17551069133529637
Zustand 8zu1:	0.06142080847078948	- = -	0.06109886512578364
Zustand 8zu2:	0.14985220801367427	- = -	0.1540595096580745
Zustand 8zu3:	0.07952631210306288	- = -	0.07324981473489553
Zustand 8zu4:	0.04512216680874241	- = -	0.038189304206464776
Zustand 8zu5:	0.06495789627477239	- = -	0.071792993045779
Zustand 8zu6:	0.13730678364533194	- = -	0.14013770635945028
Zustand 8zu7:	0.05050998992182979	- = -	0.05370379979050981
Zustand 8zu8:	0.1501562363919979	- = -	0.16857469908807282