

Contents

1 Dating mit Julia 1

1 Dating mit Julia

1.1 Vorwort

Bei der Auswahl des Projekts war uns schnell bewusst, dass wir etwas programmieren wollen, um die im Kurs erlernten mathematischen Theorien anzuwenden. Schon während der Vorlesung zum Thema Support Vector Machines wurde ersichtlich, welche Möglichkeiten ein Programm, das auf diesem Konzept beruht, bietet und so entschieden wir uns, ein solches Programm als Projekt zu schreiben. Als Anwendungsgebiet wählten wir die Partnersuche. Das im Folgenden vorgestellte Programm ist in der Lage, auszugeben, ob eine Person eine andere Person äußerlich anspricht, sofern diese Person zuvor Porträtbilder bewertet hat.

1.2 Beschreibung

Das Program ist so aufgebaut, dass eine Testperson zuerst über zweihundert Personen mit "hot" oder "not" bewertet. Dabei erfahren wir etwas über die Ansprüche beziehungsweise Vorlieben der Person. Anschließend wird eine neue Person dem Program hinzugefügt. Die Merkmale dieser neuen Person werden mit denen der Trainingsdaten verglichen. Im Anschluss gibt der Computer aus, ob diese neue Person den Vorlieben der Testperson entspricht.

1.3 Datensatz

Als Erstes benötigten wir einen großen Datensatz, also eine große Menge an Bildern, die später zum Trainieren des Programs auf die Vorlieben einer Person genutzt werden sollten. Hierbei gilt, dass ein grßer Datensatz zu besseren Ergebnissen fhrt. Da das Program für ein Klientel von Frauen im Alter von etwa 15 – 25 Jahren geschrieben werden sollte, suchten wir 212 Bilder von Mern im Alter von 15 - 50 Jahren aus, wobei die gro Mehrheit sich im Bereich zwischen 20 und 30 Jahren befindet. Bei der Auswahl wurde auf einen möglichst realistischen Datensatz geachtet. So sind Personen aus allen Kontinenten vertreten, mehr Braunhaarige als Blonde, mehr Blonde als Rothaarige, und so weiter. Diese Bilder wurden nach 13 Features klassifiziert. Diese sind das Alter, die Ethnie, die Gesichtsform, die Hautfarbe, die Haarfarbe, die Frisur, die Gesichtsbehaarung, die Augenbrauen, die Augenfarbe, die Lippen, die Nase und die eventuell vorhandene Brille und der Körperschmuck.

1.4 Featurization

Fr die Featurization, also das Eintragen von Features, haben wir Einheitsvektoren verwendet, damit kein Feature bevorzugt wird. Das hei, dass jeder Vektor die Le Eins besitzt und auch die Abste zwischen einem Paar von Vektoren betr Eins. Die gleichen Abste bedingen, dass jeder Vektor gleich behandelt wird. Die Vektoren unterscheiden sich nur dadurch, an welcher Komponente der Wert ungleich null ist. Bei einem trifft die Eigenschaft zu, der Rest des Abschnitts (Untervektoren) wird durch Nullen aufgefüllt. Man kann auch zwei Ausprngen einer Eigenschaft aufzeigen, indem man beispielsweise 0.5 und 0.5 nimmt. Wichtig ist hierbei, dass immer Eins als Gesamtbetrag des Abschnittes/Untervektors heraus kommt.

1.5 Program

Die Werte der Trainingsdaten speichern wir in einer Matrix ab und lassen sie dann von unserer Support Vector Machine klassifizieren. Es sucht also nach Verbindungen zwischen den Eigenschaften der Bilder. Es wird geschaut, ob Zusammenhänge zwischen allen positiv bewerteten Person existieren. Auf Grund dieser Auswertung erstellt der Computer eine Funktion, die die verschiedenen Eigenschaften, wie beispielsweise Haarfarbe, den Ausgabewerten "Hot" Und "Not" zuweist.