Intelligente Datenanalyse in Matlab

Prüfung: Projektaufgabe 2

Paul Prasse
Christoph Sawade
Prof. Tobias Scheffer

Datum: 01.01.2012

Die Projektaufgabe ist Teil der Prüfung *Intelligente Datenanalyse in Matlab*. Jede Aufgabe soll durch einen Studenten selbstständig bearbeitet und die Lösung innerhalb der mündlichen Prüfung vorgestellt werden. Ein Ausdruck des MATLAB-Programmcodes und der Ergebnisse in Form eines Diagramms, Tabelle o.ä. werden vorausgesetzt; die Art der Präsentation der Ergebnisse ist dem Studenten freigestellt.

Problemstellung

Auf einem Bewertungsportal können Nutzer Beiträge zu Produkten, Dienstleistungen usw. verfassen. Die Qualität dieser Beiträge ist jedoch stark unterschiedlich. Zu diesem Zweck können andere Nutzer die Beiträge auf einer Skala von 1 bis 5 Sternen bewerten. Der Betreiber des Bewertungsportals möchte diese Information nutzen um jeweils nur die drei höchst bewerteten Beiträge anzuzeigen. Allerdings sind auch die Bewertungen der Beiträge höchst subjektiv.

Sie wurden damit beauftragt ein Verfahren zu implementieren welches basierend auf den Nutzer-Bewertungen die Qualität der Beiträge bestimmt. Dafür wurden für alle 1.000 Nutzer des Portals die abgegebenen Bewertungen gesammelt und im Format „Bewertender Nutzer“, „Nutzer welcher den Beitrag verfasst hat“ und „Bewertung“ gespeichert. Für jedes Produkt existieren Beiträge unterschiedlicher Nutzer. Ihr Modell soll die Frage beantworten welche drei Beiträge davon angezeigt werden sollen.

Aufgabe

Importieren sie die Daten in MATLAB und wählen sie eine geeignete Repräsentation der Daten. Die Eingabe des Modells ist ein binärer Vektor (für jedes Produkt) der angibt welcher Nutzer einen Beitrag geschrieben hat. Die Ausgabe sollen die Indizes der besten drei Beiträge bzw. deren Autoren sein. Identifizieren sie ein Verfahren für dieses Lernproblem, implementieren sie das Verfahren in MATLAB und trainieren sie ein Modell. Überlegen sie wie das Verfahren evaluiert werden kann. Begründen und dokumentieren sie kurz alle durchgeführten Schritte.