

**Projektbericht zum Modul Information Retrieval und
Visualisierung Sommersemester 2021**

Visualisierung von Daten des Videospiels Fifa 19

Johannes Lange

13. August 2021

1 Einleitung

Tamara Munzner angucken

Zielproblem formulieren und daraus Fragestellung schreiben -> Wie können diese mittels Visualisierung beantwortet werden

Die Visualisierung von Daten nimmt mit Hinblick auf Big Data und damit immer unübersichtlicheren Grunddaten an Bedeutung zu. Um aus großen Mengen von Daten neue Informationen zu gewinnen reicht es nicht aus die Daten direkt zu analysieren. Durch die Anwendung der richtigen Visualisierungstechniken können neue Informationen gewonnen werden. In diesem Bericht geht es um die Visualisierung von Daten aus dem Computerspiel Fifa 19. Da sich diese Daten allerdings auf tatsächliche Fussballer beziehen besteht die Hoffnung mittels Visualisierung dieser Daten Rückschlüsse auf die tatsächlichen Sportler schließen zu können.

1.1 Anwendungshintergrund

Was sind die Daten, sind die Visualisierungslösungen angebracht, möglicherweise auf einzelne Variablen eingehen[1]

Die Frage, welche sich bei diesen Daten stellt ist, welchen Nutzen sie außerhalb eines Videospiels haben. Deswegen ist mein Ansatz zu vergleichen wie genau diese Daten die Wirklichkeit widerspiegeln. Ein Beispiel dafür könnte sein ob sich das Rating eines Spielers mit der körperlichen Entwicklung eines Spielers bewegt. Sollte dies zutreffen müssten Spieler im Alter ihres physischen Peaks das höchste Rating haben.

1.2 Zielgruppen

Fifa 19 Spieler -> Viel Vorwissen, Fussball Fan -> recht viel Vorwissen, Sportinteressent -> Kaum Vorwissen

Zielgruppen für Scatterplot:

Parallele Koordinaten:

Baumdigramm: Hier lässt sich erkennen welche Sportliga Europas im Durchschnitt die besten Spieler hat. Deswegen ist dies für X interessant.

Mögliche Zielgruppen: Videospieler, Teilnehmer einer Fantasy Fußball Liga

1.3 Überblick und Beiträge

Sehr kurze Beschreibung der Daten, der Drei Techniken und Beiträge des Projekts(Mehrwert der Visualisierungstechniken)

Als erste Visualisierung habe ich mich für einen Scatterplot entschieden um mit diesem einen groben Überblick über mögliche Trends in den Daten zu ermöglichen.

Als zweite Technik habe mich für die Technik der parallelen Koordinaten entschieden, mit dieser lassen sich Zusammenhänge zwischen zwei Merkmalen gut erkennen.

Als dritte und letzte Technik habe ich die Baumdarstellung ausgewählt um sichtbar zu machen welche Fussballliga Europas im Durchschnitt die besten Spieler hat. Deswegen ist dies für X interessant.

2 Daten

Beschreibung der gegebenen Daten, Sind sie zur Beantwortung der Fragestellungen geeignet? Welche zusätzlichen Daten wurden genutzt

Grundsätzlich eignen sich die Daten gut um die gewünschten Fragestellungen beantworten zu können, jedoch enthält der Grunddatensatz einige Felder, die für die Visualisierung nicht nötig sind, deswegen wurde der Datensatz in der Vorvorarbeitung noch verkleinert (Siehe 2.2 auf Seite 3). Außerdem ist der Datensatz sehr groß, was gerade bei den parallelen Koordinaten zu Problemen führen kann wenn der ganze Datensatz angezeigt wird, deswegen wurde sich bei den parallelen Koordinaten dazu entschieden nach zusätzlichen Dimensionen wie Nationalität zu Filtern um dies so übersichtlicher zu gestalten. Da Datenwerte wie Größe, Alter und Rating der Spieler diskret sind wurde sich dazu entschieden im Scatterplot die Anzahl an Spielern welche in diesem Punkt enthalten sind auszugeben. Weiterhin wurde die Opazität der Punkte verringert, da so zu sehen ist an welchen Stellen sich mehrere Spieler überlagern.

2.1 Technische Bereitstellung der Daten

Über Github gehostet, als CSV für 1 und 2 und als JSON für 3. Die Daten mussten im Programm geparsed werden.

2.2 Datenvorverarbeitung

- Daten mussten gekürzt werden, da viele Visualisierungstechniken sonst unübersichtlich wurden.
- Bei 3 wurde sich auf einzelne Fussballligen konzentriert -Weiterhin sind einige Spalten der Daten sehr irrelevant für die Analyse (Beispiel: Positionen der Spieler)

3 Visualisierungen

3.1 Analyse der Anwendungsaufgaben

Wie hilft Scatterplot/Parallele Koordinaten/Baumdiagramm die genannten Problemstellungen zu beantworten?

3.2 Anforderungen an die Visualisierungen

Wie muss die Visualisierung designed werden um das Zielproblem gut beantworten zu können?

3.3 Präsentation der Visualisierungen

Visualisierungstechniken vorstellen, Interaktivität zeigen, Designentscheidungen begründen(Erfüllen diese die Anforderungen?), Diskutieren wieso nicht andere Techniken verwendet wurden(Expressivität und Effektivität).

3.3.1 Visualisierung Eins

3.3.2 Visualisierung Zwei

3.3.3 Visualisierung Drei

3.4 Interaktion

Interaktionen in den Visualisierungen(möglicherweise Interaktion zwischen den Techniken), Warum genau diese Techniken, welche Zwecke erfüllen sie für die Anwender, Warum wurden andere nicht umgesetzt

4 Implementierung

Wie ist der Quellcode gegliedert, was lies sich aus den Übungen übernehmen, Wie sieht die Datenstruktur des Modells aus -> in dem verschiedene Zustände der Interaktion gespeichert wurden (Success record)

5 Anwendungsfälle

Spezifischen Anwendungsfall für Nutzergruppen vorstellen, der an Hand der Visualisierungstechniken visuell erkennbar ist, wäre dies auch mit anderen Techniken möglich gewesen? Aufwand mit anderen Techniken vergleichen

5.1 Anwendung Visualisierung Eins

5.2 Anwendung Visualisierung Zwei

5.3 Anwendung Visualisierung Drei

6 Verwandte Arbeiten

Zwei Artikel mit ähnlichen Zielen diskutieren

7 Zusammenfassung und Ausblick

Beiträge der Anwendung, welcher Mehrwert für Zielgruppe entsteht, mögliche Erweiterungen(Visualisierungen oder Daten)

Anhang: Git-Historie

Literatur

- [1] Tamara Munzner. “Process and Pitfalls in Writing Information Visualization Research Papers”. In: *Information Visualization: Human-Centered Issues and Perspectives*. Hrsg. von Andreas Kerren u. a. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, S. 134–153. ISBN: 978-3-540-70956-5. DOI: 10.1007/978-3-540-70956-5_6. URL: https://doi.org/10.1007/978-3-540-70956-5_6.