

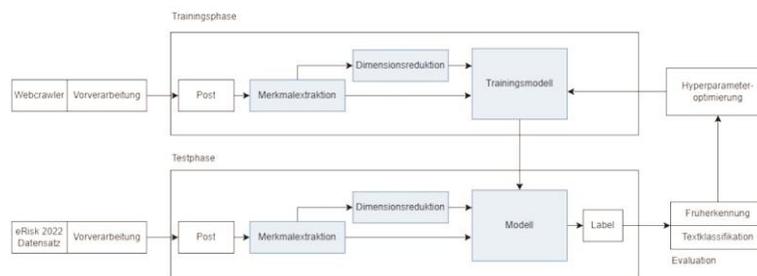
## Der Einsatz ist hoch - Erkennung von Spielsucht mittels Information Retrieval und Machine Learning

Das Ziel dieser Arbeit ist es, ein System zu entwickeln, welches die Erkennung von Spielsüchtigen in Foren verbessert. Menschen mit Spielsucht leiden unter dem zwanghaften Drang, Glücksspiele zu spielen. Dies wirkt sich nachhaltig auf das psychische Wohlbefinden aus und beeinflusst längerfristig die Lebensumstände und das nähere Umfeld der betroffenen Person. Durch frühzeitige Erkennung gefährdeter Menschen können diese informiert und zu entsprechenden Massnahmen bewegt werden. Das System der Textklassifikation basiert auf Ansätzen und Ideen, welche 2021 im Rahmen verschiedener Arbeiten des CLEF eRisk Lab angewendet wurden. Für jeden geschriebenen Post eines Users in einem Forum wird beurteilt, ob die Person potenziell spielsüchtig ist oder nicht. Für diese Beurteilung wurden Anpassungen an der Merkmalsextraktion und der Klassifikation vorgenommen. Für den Trainingsdatensatz wurden die Daten von Reddit Posts mittels eines Webcrawlers gesammelt. Die Ergebnisse der eRisk 2022 zeigen, dass XGBoost leicht besser als SVM abschneidet. Für eRisk 2021 ist das alte System von Team UNSL mit BoW und SVM immer noch am stärksten. Einzig der TPR-Wert ist mit GloVe und XGBoost optimaler als das Referenzsystem.



Diplomierende  
Samuel Stalder  
Erman Zankov

Dozent  
Martin Braschler



Optimierungszyklus der Ansätze mit  
maschinellern Lernen