

Abstract zur Dissertation „Informationen zur Bonitätsprüfung auf Basis von Daten aus sozialen Medien“

Die Dissertation „Informationen zur Bonitätsprüfung auf Basis von Daten aus sozialen Medien“ befasst sich mit der Frage, inwiefern sich in Daten aus sozialen Medien Hinweise finden lassen, welche die Informationsbasis von Bonitätsprüfungen ergänzen können. Nach einer Einleitung, in welcher die Defizite gegenwärtiger Bonitätsprüfungsverfahren aufgezeigt werden, und den anschließend dargelegten Grundlagen werden der aktuelle Stand in Forschung und Praxis des Themas aufgearbeitet. Hierbei wird deutlich, dass insbesondere der Nutzen von Textdaten aus sozialen Medien wie z. B. Blogs für Bonitätsprüfungen von Unternehmen bislang nicht untersucht wurde und diesbezüglich Forschungslücken und nicht erschlossene Märkte vorhanden sind.

In einer anschließenden Machbarkeitsstudie wird nachgewiesen, dass in Textdaten der Plattform Twitter Hinweise zu finden sind, die auf die zukünftige Insolvenz eines Unternehmens hindeuten. Weiterhin ist die Anzahl der veröffentlichten nutzergenerierten Textdaten vor einer Insolvenz signifikant höher als in einer Vergleichsgruppe. Anhand von monatlichen Stimmungsindikatoren, welche die Textdaten in negativ, neutral oder positiv im Hinblick auf die finanzielle Stabilität eines Unternehmens einteilen, können zukünftig insolvente Unternehmen von solventen unterschieden werden.

Daraufhin wird mit Hilfe von Laborexperimenten analysiert, welche Verfahren sich am besten eignen, um die in der Machbarkeitsstudie manuell klassifizierten Textdaten automatisiert auszuwerten. Es wird deutlich, dass maschinelle Lernalgorithmen den technisch einfacher anzuwendenden Stimmungswörterbüchern überlegen sind. Weiterhin wird aufgezeigt, inwiefern die Textdaten vorverarbeitet werden sollten, um bestmögliche Analyseergebnisse zu generieren.

Im letzten Schritt wird ein Entscheidungsunterstützungssystem, welches Kreditentscheider nutzen können, um Textdaten aus sozialen Medien im Rahmen von Unternehmensbonitätsprüfungen zu verwenden, konzipiert, implementiert und evaluiert. Im Rahmen der Systemkonzeption werden funktionale Anforderungen und Qualitätsanforderungen an die Benutzungsoberfläche, die Anwendungsebene und die Datenbank eines solchen Systems sowie Randbedingungen erhoben, woraufhin die Modellierung der Ebenen erfolgt. Nach der Implementierung wird die Evaluation unter Zuhilfenahme von Experten aus der Praxis fachlich durchgeführt. Die technische Evaluation wird mit Hilfe von Anwendungstests verifiziert. Während der Evaluation ergeben sich weitere Anforderungen. Das Ergebnis ist ein Leitfaden, der es Wissenschaftlern und Praktikern ermöglicht, eine Instanz eines solchen Systems zu entwickeln.

Die Arbeit schließt mit einem Fazit, in welchem die Ergebnisse zusammengefasst, Limitationen aufgezeigt und ein Ausblick auf zukünftige Forschungsvorhaben gegeben wird. Neben der Analyse von weiteren heterogenen Textdatenbeständen ist eine Instanz des entwickelten Systems in einem Feldexperiment zu evaluieren, um Auswirkungen auf Bonitätsprüfungsprozesse zu analysieren.

Abstract for the Dissertation „Information for Credit Risk Assessment Based on Data from Social Media Platforms“

The dissertation „Information for Credit Risk Assessment Based on Data from Social Media Platforms“ answers the question, in how far hints, which can be used to improve the information base of credit ratings, can be found in social mediums. After in introduction, in which deficits of current credit risk assessment procedures are highlighted, and an elucidation of the foundations, the state of the art in research and practice are reconditioned. It becomes evident, that especially the use of textual data from social mediums e.g. blogs für corporate credit risk assessment has not been investigated yet and hence, research gaps and unexploited markets are present.

In a following feasibility study it is demonstrated, that hints for an impending insolvency of a company can be found in textual data from the platform Twitter. Furthermore, the number of published user generated texts before an insolvency is significantly higher than for a solvent reference group. By use of monthly aggregated sentiment scores, which classify texts as negativ, neutral or positive regarding the financial stability of a company, future insolvent companies can be distinguished from solvent counterparts.

Thereupon, it is analysed by laboratory experiments, which procedures are best suited to automatically interpret the textual data which has been manually coded during the feasibility study. It becomes evident, that machine learning algorithms outperform sentiment dictionaries, which are less complex. Furthermore, it is shown how the textual data should be preprocessed in order to attain best possible classification results.

In the last step, a decision support system, which can be used by credit decision makers in order to use textual data from social media platforms when analyzing the credit worthiness of companies, is conceptualized, implemented and evaluated. In the course of the conceptualization, functional and qualitative requirements regarding the user interface, the application layer and the database as well as boundary conditions are compiled, whereupon the modelling of the three layers is conducted. After the implementation, the evaluation is performed by utilization of experts from practice. A technical evlauation is undertaken by use tests. During the evaluation, additional requirements have been revealed. The result is a guideline which enables scientists and practioners to develop an instance of the proposed system.

The dissertation ends with a conclusion in which the results are summarized, limitations are highlighted and an outlook to future research endeavors is given. Next to the analysis of further, heterogeneous textual data, an instance of the developed system should be evaluated within a field experiment in order to analyze effects on credit risk assessment processes.