

Klassifikation regulatorischer Pressemitteilungen im Themenfeld Recovery und Resolution

Studiengang: MAS Data Science

Die Auswertung regulatorischer Pressemitteilungen für das Thema Recovery und Resolution erfolgt in der FINMA bislang manuell durch Fachmitarbeitende. Intelligente Lösungen, die eine automatisierte Verarbeitung der Texte ermöglichen, fehlen. Diese explorative Masterarbeit untersucht verschiedene Textklassifikationsverfahren und evaluiert sie im Hinblick auf Klassifikationsleistung und praktischen Nutzen.

Ausgangslage

Aufsichtsbehörden sowie Institutionen in der Finanzindustrie veröffentlichen regelmässig Pressemitteilungen zu vielfältigen Themen. Für das Themenfeld Recovery und Resolution sind diese zentral, da sie Transparenz schaffen und Hinweise auf regulatorische Entwicklungen zur Stabilisierungs- oder Notfallplanung systemrelevanter Institute geben. Bislang erfolgt die Identifikation relevanter Mitteilungen manuell – ein zeitintensives Vorgehen mit dem Risiko, Entwicklungen verspätet wahrzunehmen. Mit der steigenden Frequenz und Heterogenität regulatorischer Pressemitteilungen wächst daher der Bedarf an automatisierten Verfahren. Methoden der natürlichen Sprachverarbeitung bieten hierfür ein hohes Potenzial. Da jedoch zahlreiche Ansätze existieren, lässt sich nicht pauschal bestimmen, welches Verfahren für den Anwendungsfall der FINMA am besten geeignet ist.

Zielsetzung

Ziel der Arbeit ist es, verschiedene Textklassifikationsverfahren auf ihre Eignung zur Identifikation regulatorischer Pressemitteilungen im Themenfeld Recovery und Resolution zu evaluieren. Untersucht werden drei methodisch unterschiedliche Ansätze: ein klassisches Verfahren (logistische Regression), ein State-of-the-Art-Ansatz mit Fine-Tuning vortrainierter BERT-Modelle sowie ein trainingsfreier Ansatz mit generativen Sprachmodellen (In-Context Learning). Im Mittelpunkt stehen sowohl die Klassifikationsleistung als auch der praktische Nutzen im FINMA-Kontext. Auf dieser Grundlage sollen Stärken und Einschränkungen aufgezeigt und konkrete Handlungsempfehlungen abgeleitet werden.

Vorgehen

Die Untersuchung basiert auf einem explorativen Forschungsdesign. Grundlage bildet ein eigens erstellter Datensatz mit 2'826 Pressemitteilungen von 13 Aufsichtsbehörden und Institutionen. In kontrollierten Experimenten wurden die verschiedenen Verfahren

umgesetzt und quantitativ evaluiert. Darauf aufbauend erfolgte eine Nutzwertanalyse, welche die Verfahren im FINMA-Kontext qualitativ bewertete. Die Kombination beider Ansätze ermöglicht eine differenzierte Einschätzung von Stärken und Einschränkungen.

Ergebnisse

Alle Verfahren erzielten eine hohe Klassifikationsleistung (Macro-F1 $\geq 0,9$), setzten jedoch unterschiedliche Schwerpunkte bei Präzision und Sensitivität. In-Context Learning überzeugte durch Flexibilität und den Verzicht auf Trainingsdaten, war jedoch mit einer Abhängigkeit von proprietären Anbietern verbunden. Das klassische Verfahren der logistischen Regression punktete durch Transparenz und geringen Rechenbedarf, war jedoch auf annotierte Daten angewiesen. Das Fine-Tuning vortrainierter BERT-Modelle konnte zwar hohe Leistungen erzielen, ging jedoch mit erhöhter Komplexität und hohem Ressourcenbedarf einher.

Als Handlungsempfehlungen sollte die FINMA eine Strategie entwickeln, die den Workflow für die Evaluation von KI-Systemen verbindlich beschreibt und so einen strukturierten, wiederholbaren Prozess schafft. Zweitens ist im untersuchten Anwendungsfall festzulegen, ob Präzision, Sensitivität oder ein ausgewogenes Verhältnis beider Grössen prioritär ist, um darauf aufbauend verbindliche Akzeptanzkriterien zu definieren. Für die Modellauswahl bietet In-Context Learning das grösste Potenzial für einen prototypischen Einsatz, während das klassische Verfahren als transparente Baseline oder Fallback-Verfahren geeignet bleibt. Fine-Tuning vortrainierter BERT-Modelle empfiehlt sich nur bei klarer Ressourcenzuteilung und dem Ziel maximaler Leistungsoptimierung.



Silvio Käser
kaeser.silvio@gmail.com