DeepScan: Maschinelles Lernen zur automatisierten Erkennung von IKT-Sicherheitsvorfällen und Manipulationsversuchen

Förderkennzeichen: 01IS18045A

1. Aufgabenstellung, wissenschaftlicher und technischer Stand

Aufgabenstellung: Während die Digitalisierung von Geschäftsabläufen zu einer steigenden Adaption von betriebswirtschaftlicher Software, insbesondere ERP-Systemen führt, steigen auch die kriminellen Aktivitäten und Betrugsfälle im digitalen Kontext stetig weiter an.
Ziel des Projekts DeepScan war es, effiziente, robuste und nachvollziehbare Machine Learning und Deep Learning Verfahren zu entwickeln und in Form einer entsprechenden Software/Toolbox bereitzustellen, welche sich einfach an Systeme anbinden lässt. Das Projekt sollte dabei sowohl Datenschutz- und IT-Sicherheitskonzepte entwickeln als auch ein Testbed aufbauen, welches die Entwicklung und die Evaluation der fokussierten Verfahren ermöglicht. Zur Entwicklung sollten mehrere Verfahren wie Metric Learning, HypTrails, SubTrails, Machine- und Deep Learning auf ihre Eignung hin untersucht werden.
Hinsichtlich der abgeleiteten Anforderungen an ein entsprechendes Anomalieerkennungsystem sollten die Methoden evaluiert und schließlich in einem Demonstrator praktisch umgesetzt werden.

Stand der Forschung: Algorithmische Verfahren zur Betrugserkennung in betriebswirtschaftlicher Software bewegten sich bisher in den Bereichen Metric-, Machine-, Deep Learning und Process Mining. Dabei lag jedoch bisher kein besonderer Fokus auf der effizienten, robusten und nachvollziehbaren Ausgestaltung der Verfahren. Mit der Entwicklung von tiefen neuronalen Netzen wurden zwar auch Architekturen zur Anomalieerkennung in anderen Domänen entwickelt, allerdings konnte die Nachvollziehbarkeit im Vergleich zu beispielsweise auf Entscheidungsbäumen basierenden oder anderen linearen Verfahren nicht gewährleistet werden.

Technischer Stand: Die aktuellen state-of-the-art Verfahren zur Betrugserkennung wie beispielsweise Audit-Trails, Reports, forensische (z.B. auditbee oder audicon) oder algorithmische Verfahren lassen sich in externe bzw. in IT-Systemen implementierte Verfahren, außerdem in Echtzeit- und zeitversetzte, sowie in holistische und atomistische Ansätze unterteilen. Aktuelle Implementierungen in ERP-Systemen beschränken sich auf nicht datengetriebene Formen, wie Audit-Trails und Reports. Datengetriebene Verfahren werden hingegen nur im Kontext einer forensischen Datenanalyse, meist erst nach einem bewussten Verdacht, eingesetzt, was unter anderem den nicht vorhandenen Datensätzen zur Entwicklung bzw. dem Training entsprechender Verfahren geschuldet ist. Echtzeit basierte Verfahren kommen im betriebswirtschaftlichen Kontext bisher kaum zum Einsatz.

2. Ablauf des Vorhabens

Zu Beginn des Vorhabens wurde zunächst in Zusammenarbeit mit datenschutz süd GmbH eine initiale Konzeption zum Umgang mit unternehmenskritischen und personenbezogenen Daten ausgearbeitet. Das daraus resultierende Daten- und IT-Sicherheitskonzept berücksichtigt die rechtlichen Anforderungen der Machine-Learning-Toolbox und stellt sicher, dass die Verarbeitung von Geschäftsdaten im ERP-System im Projekt mit den geltenden Datenschutzbestimmungen vereinbar ist. Daneben wurde mit der Untersuchung von Datenstrukturen von ERP-Systemen und durch die Bereitstellung eines Testbeds eine wesentliche Grundlage für das Projekt geschaffen, um die im Projektverlauf betrachteten Verfahren zur Modellierung von Vorgängen und Erkennung von Manipulationsversuchen in ERP-Systemen zu entwickeln und evaluieren. In diesem Testbed wurden

basierend auf verschiedenen Verfahren zur Datengenerierung diverse synthetische Datensätze unterschiedlicher Struktur und Güte erzeugt, die die Basis für die weiteren Experimente darstellen und anderen Forschern öffentlich zur Verfügung gestellt wurden.
Auf Basis einer strukturierten Anforderungsanalyse für die Modellierung von Vorgängen und Prozessen in ERP Systemen und Vorarbeiten aus der Literatur, wie beispielsweise Word2Vec, wurden erfolgreich semantische Einbettungen von Abläufen und Daten in ERP Systemen erzeugt und verschiedene Businessszenarien auf einer hoher Abstraktionsebene hypothesenbasiert modelliert, was die Anwendung von HypTrails ermöglichte.
Neben der Untersuchung von Distanzmaßen, Metric Learning und hypothesen-basierten Ansätzen konnten bei der Betrugserkennung vor allem mit End-to-end Verfahren hervorragende Ergebnisse erzielt werden. Hierfür wurde eigens die neuronale Netz-Architektur iNALU zur Modellierung von numerischen Zusammenhängen entwickelt und zur Anwendung gebracht, deren Vorhersagen durch SHAP-basierte post-hoc Methoden erklärbar gemacht werden konnten. Mittels einer im Vorhaben entwickelten Evaluationsmethodik der Erklärbarkeit wurden verschiedene Modelle untersucht und solche identifiziert, die sowohl gute Erkennungsleistung als auch gute Nachvollziehbarkeit für den Einsatz in der Praxis bieten, was in mehreren Publikationen veröffentlicht wurde.
Diese vorwiegend auf kleinen Testdatensätzen aus dem Testbed entwickelten Methoden wurden schließlich hinsichtlich ihrer Dynamik und Effizient auf einem Big-Data Realdatensatz optimiert sowie die Robustheit der Methode gegenüber variierender Datencharakteristika untersucht. Zur Vorbereitung einer offline Evaluation der Methodik auf großen Datensätzen aus der Praxis wurde des Weiteren ein Annotationstool entwickelt und ein modulares Active-Learning Framework konzipiert. Zuletzt wurden die im Projekt als vielversprechend bewerteten Methoden für die Ausgestaltung einer Machine Learning Toolbox als ERP-Add-in vorbereitet und basierend auf den Anforderungen, die mit dem Projektpartner Step Ahead sowie dem Mitglied des Projektbeirats weclapp für ein Add-in der jeweiligen ERP Systeme identifiziert werden konnten, ein cloud-fähiger Microservice zur Kapselung der im Projekt entwickelten Verfahren zur Anomaliedetektion entwickelt. Die Evaluation der Komponenten wurde quantitativ und qualitativ durchgeführt, wodurch nicht nur für simulierte Angriffe, sondern auch auf Echtdaten erste positive Ergebnisse gezeigt werden konnten.

3. Wesentliche Ergebnisse und Zusammenarbeiten

Aus dem Forschungsvorhaben gingen mehrere wissenschaftliche Publikationen zu den unterschiedlichen Zielsetzungen des Projekts hervor. Sowohl für die Entwicklung von Algorithmen zur Manipulationserkennung, als auch für die umfassende Auswertung der Erklärbarkeit von Manipulationserkennungsmethoden wurden Forschungsbeiträge auf internationalen Konferenzen veröffentlicht. Zusätzlich wurden zur Generierung von ERP System Daten mit enthaltenen Manipulationen mehrere Beiträge publiziert und daraus resultierende Datensätze der Öffentlichkeit zur Verfügung gestellt. Auch die Konzipierung von Manipulationserkennungssystemen auf ERP System Daten in der Praxis wurde formalisiert und als Forschungsbeitrag veröffentlicht. Neben diesen wissenschaftlichen Resultaten wurden im Projekt weitere Ergebnisse bei der Formalisierung eines Datenschutz- und IT-Sicherheitskonzepts erzielt, sowie erfolgreich ein ERP Add-in zur Manipulationserkennung im Produktivsystem erstellt. Projektübergreifende Kooperationen bestanden insbesondere durch den Austausch mit Mitarbeitern der Hochschule Coburg bei der Entwicklung von Methoden zur Manipulationserkennung sowie durch den Austausch mit dem Softwareunternehmen Celonis bei der Modellierung von ERP System Daten. Zusätzlich bestanden Kooperationen mit Mitarbeitern der Hochschule Karlsruhe, welche Ergebnisse unseres Projekts im KOEX Forschungsprojekt nachnutzen wollen.

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

PROJEKT DEEPSCAN

Maschinelles Lernen zur automatisierten Erkennung
von IKT-Sicherheitsvorfällen und Manipulationsversuchen

Fachlicher Sachbericht Teil II: Eingehende Darstellung

Gefördert vom

Bundesministerium
für Bildung
und Forschung

Betreut vom

DLR Projektträger

Arbeitspaket A: Projektmanagement

Zusammenfassung:

Das Arbeitspaket A umfasst die wesentlichen Tätigkeiten des Projektmanagements. Im Folgenden wird ein Kurzüberblick über die Planung und Steuerung des Konsortialprojekts gegeben. Dabei wird maßgeblich auf den Projektplan, die Meilensteine und relevante Anpassungen eingegangen.

Inhaltsverzeichnis Arbeitspaket A

1 Projektplanung

Die Projektplanung des Vorhabens fand basierend auf den im Langantrag bereits spezifizierten Arbeitspaketen AP A - H statt. Neben den spezifischen Projektinhalten wurde zum Auftakt der Arbeitspakete jeweils eine entsprechende Planungsrunde durchgeführt, um etwaige Anpassungen aufgrund der Ergebnisse vorangegangener Arbeitspakete in eine aktualisierte Planung zu überführen. Eine Übersicht der Planung kann Abb. 1 entnommen werden.

2 Verwendete Werkzeuge

Zur Planung wurde die Software OpenProject eingesetzt, welche neben der einfachen übersichtlichen Planung auch ein entsprechendes Projektcontrolling über alle Parteien ermöglichte. Neben den Arbeitspaketen wurden in OpenProject auch Zwischen- und Abschlussberichte koordiniert und Planungsanpassungen durchgeführt.

Für die Koordination der Entwicklungen wurde das Uni-interne GIT-Lab verwendet, publizierte Inhalte wurden, um diese zugänglich zu machen, in Github-Repositories überführt.

3 Projektdurchführung

Während der Projektdurchführung gab es diverse Einflüsse, die den geplanten Projektverlauf verändert haben.
Größter Einfluss auf das Projekt hatte dabei die Coronapandemie. Aufgrund der anfänglichen Unsicherheiten, Lockdowns und diversen Regelungen

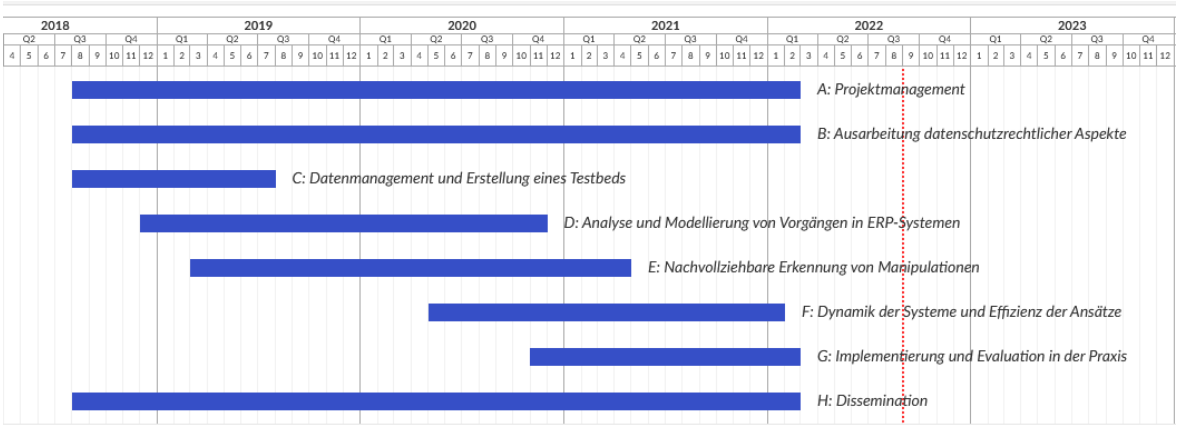

Abbildung 1: Projektplanung

von Firmen, mussten bspw. viele der in den Arbeitspaketen geplanten Anforderungsanalysen auf virtuellem Weg durchgeführt werden. Obwohl es zu keinen personellen Engpässen kam, führten einzelne Krankheitsfälle zu Verschiebungen. Genauso wurden Projektkonsortialmeetings teilweise, Corona-bedingt, digital durchgeführt und auch wissenschaftliche Konferenzen kamen teilweise nicht zustande bzw. wurden virtuell gehalten.

Neben der Corona-Pandemie können als wesentliche Einflussfaktoren außerdem die personellen und betrieblichen Umstrukturierungen genannt werden. Die Umfirmierung der Godesys zur Step Ahead GmbH führte dabei zu einigen personellen, aber auch inhaltlichen Anpassungen. Da die Step Ahead strategisch STEPS ERP als Basis für zukünftige Neukunden fokussiert, wurde im Projekt STEPS ERP als Grundlage für die Entwicklung genutzt, welches durch den überarbeiteten Aufbau und die vorhandenen Schnittstellen eine Vereinfachung der Implementation und auch eine moderne Architektur der ERP-Add-Ins ermöglichte (vgl. AP C).

Beide Einflussfaktoren, Pandemie und strukturelle Veränderungen des Konsortiums konnten durch eine Anpassung der Planung und entsprechende inhaltliche Anpassungen ausgeglichen werden. Die wesentlichste Anpassung war dabei die Verlängerung des Projekts um fünf Monate bis Februar 2022.

4 Projektmeilensteine

Die geplanten Projektmeilensteine M1-M10 wurden zur Vorbereitung der jeweiligen Arbeitspakete detaillierter ausgestaltet, inhaltlich, aber beibehalten. Tabelle 1 gibt eine Übersicht über die Projektmeilensteine. Insgesamt wurden alle Projektmeilensteine erreicht. Abbruchbedingungen, die das Projektergebnis soweit negativ beeinflusst hätten, dass eine Weiterführung nicht möglich wäre, traten demnach keine ein.

Tabelle 1: Übersicht Projektmeilensteine

#	Beschreibung	Erreichung
M1	Datenschutz Implementierungskonzepts	erfüllt
M2	Fertigstellung des Anforderungskonzepts für die Ausgestaltung der Machine Learning Algorithmen	erfüllt
M3	Fertigstellung des benötigten Testbed	erfüllt
M4	Fertigstellung des Anforderungskonzepts für die Ausgestaltung der ERP Ablauf Strukturierung und Modellierung	erfüllt
M5	Formalisierte Modellierung der relevanten Komponenten des Anwendungs-ERP Systems der godesys AG	erfüllt
M6	Fertigstellung des Anforderungskonzepts für die erklärbare Erkennung von Manipulationen in ERP Systemen	erfüllt
M7	Nutzungskonzept der ausgewählten, geeignetsten Methoden des Maschinellen Lernens	erfüllt
M8	Fertigstellung des Anforderungskonzept zur Berücksichtigung des Regelbetriebs von ERP Systemen und angemessener Dynamik	erfüllt
M9	Evaluationsergebnis eines geeigneten Active Learning Szenarios	erfüllt
M10	Projektabschluss	erfüllt

Arbeitspaket B: Ausarbeitung der datenschutzrechtlichen Anforderungen

Überblick:

Unternehmensdaten können reale Betrugsszenarien beinhalten und bieten somit den größten Mehrwert hinsichtlich Optimierung der Erkennung und Transparenz der zu entwickelnden Machine Learning Toolbox. Die Voraussetzung für den Einsatz erfordert die kontinuierliche Überwachung und Einhaltung der rechtlichen Richtlinien, insbesondere der Datenschutz-Grundverordnung (DSGVO). Zusätzlich können Daten unternehmenskritischen Charakter besitzen und müssen dementsprechend geschützt werden. Hierfür wurden Prozessstrukturen analysiert und in einem Datenschutzkonzept entsprechend berücksichtigt. Durch den generalisierten Ansatz wurde die Heterogenität von IT-Infrastrukturen berücksichtigt und diese Anforderungen in ein IT-Sicherheitskonzept überführt. Die im Rahmen des Arbeitspakets erarbeiteten Konzepte wurden im Projektverlauf begleitend in den weiteren Arbeitspaketen berücksichtigt.

Inhaltsverzeichnis Arbeitspaket B

1 Untergliederung in Teilarbeitspakete

Während die zu verarbeitende Datenmenge von Unternehmen kontinuierlich zunimmt, wächst auch das Bewusstsein, welche Art von Daten überhaupt verarbeitet werden dürfen. So gilt es insbesondere, unternehmenskritische (AP B1) und personenbezogene Daten (AP B2) zu schützen. Hinzu kommt, dass der richtige Umgang auch von der IT-Sicherheit maßgeblich beeinflusst wird. AP B3 thematisiert deswegen das im Projekt entwickelte IT-Sicherheitskonzept und AP B4 die Ausgestaltung des entwickelten Demonstrators.

2 Teilarbeitspaket B1: Konzept zum Umgang mit unternehmenskritischen Daten

Unternehmenskritische Daten repräsentieren Informationen, die für den Betrieb eines Unternehmens essenziell sind (Soliman, 2003). Aktuell existiert keine eindeutige Definition, was genau unter die Kategorie der unternehmenskritischen Daten fällt. Generell hängt die Einstufung, ob Daten unternehmenskritisch sind, von der Art und Branche eines Unternehmens ab (Soliman, 2003). Gängige Beispiele für unternehmenskritische Daten, die auf viele Branchen zutreffen, sind beispielsweise wettbewerbsrelevante Daten, juristische Informationen oder personenbezogene Daten. Da diese Daten für ein Unternehmen von besonderer Relevanz sind, gilt es, diese im Projektkontext besonders zu schützen.

Im Rahmen dieses Teilarbeitspakets wurde ein Konzept für den Umgang mit unternehmenskritischen Daten entwickelt. Hierfür wurden best-practices gesammelt und auf Basis des BSI Grundschutzes ein Maßnahmenkatalog ausgearbeitet, um unternehmenskritischen Daten einen besonderen Schutz bei deren Erhebung und Verarbeitung zukommen zu lassen. Im ersten Schritt wurde exemplarisch für unser Projektziel definiert, welche Daten unternehmenskritisch sind. Um diese adäquat zu schützen, kommen anschließend technische Maßnahmen zum Einsatz, beispielsweise die Pseudonymisierung kategorischer oder textbezogener Datenfelder oder die Anwendung spezifischer

Transformationen und Selektionen bestimmter Datenfelder. Daneben muss der Datenträger, welcher die unternehmenskritischen Daten enthält, durch bestimmte Vorkehrungen vor Manipulationen, technischen Fehlern und unberechtigtem Zugriff geschützt werden, was durch die Anwendung kryptographischer Verfahren erfolgt.

3 Teilarbeitspaket B2: Konzept zum Umgang mit personenbezogenen Daten

Um ein angemessenes Schutzniveau für die Eintrittswahrscheinlichkeiten und Schwere von Risiken für die Rechte und Freiheiten der Mitarbeiter zu gewährleisten, müssen die externen Unternehmensdaten im ersten Schritt dementsprechend vorbereitet werden. In Anlehnung an den branchenübergreifenden Standardprozess für Data Mining (CRISP-DM) wurden die datenschutzrechtlichen Anforderungen für die einzelnen Prozessschritte adaptiert. In Bezug auf die Entwicklung des Demonstrators entsteht vor allem in den ersten Prozessschritten, also Datenauswahl und -selektion der primäre Handlungsbedarf. Da die Performanz des Demonstrators neben der Wahl des Algorithmus auch den zugrundeliegenden Daten obliegt, stellen Unternehmensdaten aus der Praxis einen Zugewinn zu synthetischen Daten dar. Allerdings benötigen diese eine zusätzliche Vorbereitung bzw. Vorselektion. Bevor die Datenextraktion im partizipierenden Unternehmen bzw. aus dem ERP-System gestartet werden konnte, mussten schutzwürdige Datenfelder und Tabellen der Datenbank vorselektiert werden. Nach Artikel 9 Abs. 1 DSGVO besteht die Notwendigkeit besonders schützenswerte Daten, bereits vor bzw. während der Extraktion zu selektieren. Hierzu wurden entsprechende Tabellen der Finanzbuchhaltung bzw. des Einkaufs eines SAP-Systems (z.B. BSEG, BKPF oder EBAN), welche durch Feld, Datenelement und Kurzbeschreibung repräsentiert werden, in die Kategorien „Besonders schützenswert“, „schützenswert“ und „nicht schützenswert“ eingeteilt. Besonders schützenswerte Daten wurden direkt entfernt. Zusätzlich wurden ausschließlich

personenbezogene Tabellen, welche für das Forschungsvorhaben unerheblich sind, wie beispielsweise Stammdaten von Beschäftigten ebenfalls entfernt.

Basierend auf den Erkenntnissen und Analysen von ERP-Systemen, deren Daten und der Expertise der datenschutz süd GmbH konnte ein Konzept zum Umgang mit personenbezogenen Daten entwickelt werden, welches die Voraussetzung und Durchführbarkeit des Forschungsvorhabens bestätigt.

Die datenschutz süd GmbH hat den Prozess der Extraktion beim Datengeber geprüft und datenschutzrechtlich beurteilt. Im Rahmen der Vorselektion sind besonders schutzwürdige Datenfelder und Tabellen von der Datenextraktion, wie besonders schutzwürdige Daten nach Art. 9 Abs. 1 DSGVO oder ausschließlich personenbezogene Tabellen, falls diese für das Forschungsvorhaben unerheblich sind (beispielsweise Stammdaten von Beschäftigten), auszuschließen. soll eine Kopie der Datenbank der bereits pseudonymisierten Inhalte generiert und nach BSI-Grundschutz (OPS.1.2.3.M13) durch AES-Verschlüsselung von 256 Bit extrahiert werden. Die weitere Verarbeitung der Daten dient ausschließlich zu Forschungszwecken innerhalb der IT-Infrastruktur der beteiligten Lehrstühle. Zusätzlich müssen die Daten durch geeignete technische und organisatorische Maßnahmen vor dem Zugriff von unberechtigten Dritten gesichert werden. Abschließend sind die Weitergabe des Wissens und Informationen sowie die Veröffentlichung von Daten, die Rückschlüsse auf personenbezogene Daten zulassen und im Zuge der wissenschaftlichen Arbeit gesammelt wurden, nicht erlaubt.

Durch die prozessuale Beschreibung und Einhaltung dieser Handlungsvorschrift werden die Anforderungen der DSGVO eingehalten und die für das Forschungsvorhaben zu verarbeitenden Daten aus dem ERP-System können konform und rechtmäßig für die Ausgestaltung und Realisierung des Demonstrators eingesetzt werden.

3.1 Teilarbeitspaket B3: Erstellung eines IT-Sicherheitskonzepts

Um die Sicherstellung der Integrität und Vertraulichkeit innerhalb des Projekts zu gewährleisten, erfolgte eine Analyse und Bewertung der IT-Infrastruktur der Universität. Hierbei wurde die Bereitstellung externer Unternehmensdaten für die Entwicklung des Demonstrators fokussiert.

In Zusammenarbeit mit dem Rechenzentrum der Universität wurden die technisch-organisatorischen Maßnahmen zur IT-Sicherheit nach Art. 32 DSGVO eruiert und ein IT-Sicherheitskonzept vorbereitet, das beispielsweise eine adäquate kryptographische Verschlüsselung, Access Control Policies und Maßnahmen zur physischen Datensicherheit umfasst. Dieses IT-Sicherheitskonzept konnte schließlich in eine Checkliste umformuliert werden, die für die Datennutzung im Projekt und die Anwendung in der Praxis ein adäquates IT-Sicherheitsniveau sicherstellt und dokumentiert, bzw. Handlungsvorschläge liefert, um dieses zu erreichen.

3.2 Teilarbeitspaket B4: IT-sicherheitskonforme Ausgestaltung des Demonstrators

Nachdem die rechtliche Grundlage für die Implementierung der Machine-Learning-Toolbox hergeleitet wurde, konnten mithilfe einer qualitativen Umfrage die praxisrelevanten Gesichtspunkte erörtert werden. Insbesondere technische Aspekte hinsichtlich des Datenschutzkonzeptes wie die Verarbeitung umfassenden Datenvolumens und die

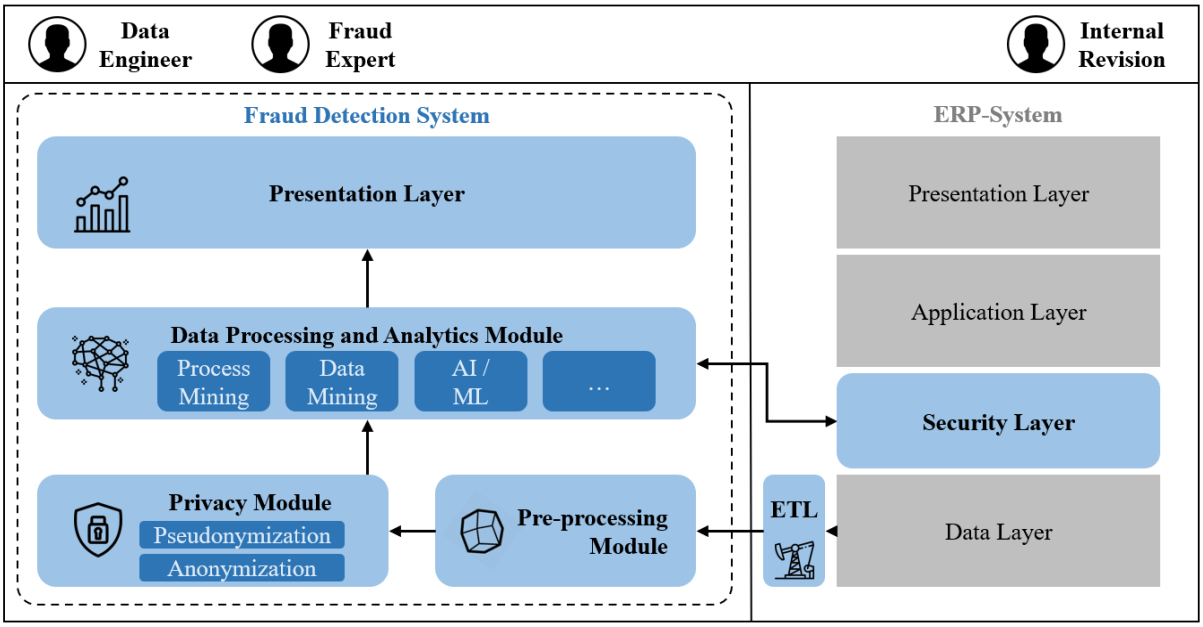

Abbildung 2: ERP-Add-In Architektur (Fuchs et al. 2021)

Integrationsfähigkeit in das ERP-System wurden thematisiert. Da der Datenaustausch über eine REST-Schnittstelle realisiert wird, werden nur die für die Untersuchung relevanten Daten abgerufen und datenschutzrelevante Informationen mittels Algorithmus anonym- bzw. pseudonymisiert. Im Rahmen des Demonstrators wurden entsprechende Maßnahmen bereits im Vorfeld bei der Datenakquise umgesetzt, um den Datentransfer konform durchzuführen. Durch die Realisierung des Demonstrators als Add-In über definierte und beschränkte Schnittstellen wird der Zugriff auf sicherheitskritische Bestandteile des übergeordneten Systems “by Design” eingeschränkt und somit eine IT-Sicherheitskonforme Ausgestaltung des Demonstrators in der Praxis realisiert.

Der Demonstrator stellt eine zusätzliche Sicherheitsschicht zwischen der Anwendungs- und der Datenschicht innerhalb des ERP-Systems dar, welche durch REST-Schnittstellen adressiert wird. Hierbei werden die von der Datenschicht erhaltenen Daten durch die Pre-Processing- und optionale Privacy-Module für die KI-Modelle vorbereitet. Durch die konforme Initialisierung des Trainingsdatensatzes können konforme und kontextfreie

Resultate erzeugt werden, welche durch die Präsentationsschicht dargestellt werden. Abbildung 1 skizziert eine Übersicht der einzelnen Abläufe und Interaktionsmöglichkeiten mit dem System und zeigt die DSGVO-konforme Realisierung des Fraud Detection Systems im ERP-System.

4 Ergebnisse des Arbeitspakets

Durch eine initiale Konzeption zum Umgang mit unternehmenskritischen und personenbezogenen Daten konnte in Zusammenarbeit mit datenschutz süd GmbH ein Daten- und IT-Sicherheitskonzept ausgearbeitet werden, welches die rechtlichen Anforderungen der Machine-Learning-Toolbox adressiert. Basierend auf den Resultaten konnte die Entwicklung des Demonstrators bzw. die Datenvorbereitung für die KI-Module datenschutzkonform durchgeführt werden.

5 Literaturverzeichnis[1]

- ***Fuchs, Anna, et al.*** *"A Meta-Model for Real-Time Fraud Detection in ERP Systems." 5 Jan. 2021, scholarspace.manoa.hawaii.edu/items/2b465ec7-18b1-4fc9-a167-e676c7ed9314/full.*

- *Soliman, F. &. (2003). The role of critical information in enterprise knowledge management. Industrial Management & Data Systems.*

[1] Eigene Publikationen wurden fettgedruckt gekennzeichnet

Arbeitspaket C: Datenmanagement und Erstellung eines Testbeds

Überblick:

Zur zielgerichteten Erfüllung des Projektziels und der Sicherstellung der Qualität wurde für das Projekt DeepScan eine ausgedehnte Anforderungsanalyse durchgeführt. Die Nutzung qualitativer Methoden ermöglicht dabei eine intersubjektive Nachprüfbarkeit anhand einer systematischen Vorgehensweise durch offengelegte Kriterien.

Neben dem klassischen Requirements Engineering wurden für das Forschungsprojekt relevante technische Aspekte und Grundlagen herausgearbeitet. Der Fokus im vorliegenden Arbeitspaket lag dabei auf den zugrunde liegenden Daten sowie den technischen Grundlagen des integrativen System-Umfelds. Im Folgenden werden die hierzu durchgeführten Arbeiten und daraus entstandenen Ergebnisse aufgearbeitet.

Inhaltsverzeichnis Arbeitspaket C

1 Methodik, Vorgehensweise und Unterteilung in Teilarbeitspakete

Ziel dieses Arbeitspakets C (Datenmanagement und Erstellung eines Testbeds) ist es, eine fundierte Grundlage für die folgenden Arbeitspakete zu schaffen. Dabei sollen die Rahmenbedingungen und Voraussetzungen für das im Projekt fokussierte Maschinelle Lernen, das Handling großer Datenmengen und die Adaptionsfähigkeit der Modelle erarbeitet werden. Arbeitspaket C unterteilt sich dazu in die Teilarbeitspakete C1 – C5.

Teilarbeitspaket C1 fokussierte die technischen Grundlagen der datentragenden Systeme. Für das Teilarbeitspaket wurden verschiedene ERP-Systeme hinsichtlich ihrer Datenerzeugung, technischen Komponenten und Erweiterbarkeit untersucht. Dabei standen insbesondere die Datenbankarten sowie die technischen Schnittstellen als Zugang zu den Daten im Fokus.
Außerdem musste in C2 sichergestellt werden, dass Betrugserkennungsansätze auch mit unvollständigen (bzw. datenschutzrechtlich eingeschränkten) Daten oder Daten aus anderen Systemen oder Systemkomponenten (Generalisierbarkeit) im weiteren Verlauf gearbeitet werden kann. In einem ersten Schritt wurden Datenbanken der ERP-Systeme untersucht und auf ihre Eignung hin analysiert. In einem zweiten Schritt wurden die von Herstellern bereitgestellten technischen Schnittstellen der Systeme analysiert.

Teilarbeitspaket C3 fokussierte die Erweiterbarkeit der Betrugserkennung auf weitere in Firmen vorhandenen Daten, also die Analyse sonstiger Unternehmensdaten in Systemen außerhalb der fokussierten ERP-Systemlandschaft. Für Teilarbeitspaket C3 wurde eine Marktanalyse durchgeführt, deren Ergebnisse dann zusammen mit den Forschungspartnern analysiert und auf ihre Eignung hin bewertet wurden.

Teilarbeitspaket C4 diente als Zwischenstands-analyse, um erste Ergebnisse des Arbeitspakets B „Datenschutz“ bei der Erstellung des Testbeds berücksichtigen zu können. De

facto gab es keine verbindlichen datenschutzrechtlichen Anforderungen, die im Rahmen des Testbeds hätten eingehalten werden müssen. Lediglich für die spätere Nutzung von Echtdaten im Rahmen der Evaluation, sowie den Produktiveinsatz der Software wurden datenschutzrechtliche Rahmenbedingungen festgehalten, diese werden aber in den assoziierten Arbeitspaketen erläutert.

Teilarbeitspaket C5 beschreibt die im Rahmen der synthetischen Datenerzeugung genutzten Methoden sowie die für das Projekt erstellte Entwicklungs- und Testumgebung. Für die Entwicklung neuer und Weiterentwicklung bestehender Machine Learning Modelle im Rahmen des Projekts, wurden möglichst heterogene Datensätze benötigt, um Adaptierbarkeit und Stabilität verschiedener Modelle und Verfahren zur Betrugserkennung testen und analysieren zu können. Dabei kamen verschiedene Verfahren zum Einsatz, um möglichst realitätsnahe synthetische Datensätze zu erzeugen. Insgesamt können die Methoden in vier verschiedene Ansätze aufgeteilt werden, welche die Datengenerierung mittels Skripten bis hin zur Datensynthese mittels neuronaler Netze umfasst. Zur Entwicklung des im Projekt fokussierten Software-Artefakts wurden die Datensätze innerhalb des Testbeds aufgearbeitet und verschiedene ML-Modellen zur Erkennung von Auffälligkeiten erprobt. Das hierzu in Teilarbeitspaket C5 entwickelte Testbed diente als Grundlage des weiteren Projektverlaufs, Die Topologie des Testbeds wird im letzten Unterkapitel detailliert beschrieben.

2 Teilarbeitspaket C1: Anforderungsanalyse

Zum Aufbau des Testbeds wurden im Teilarbeitspaket C1 die gängigen Möglichkeiten zur Extraktion von Daten, die verwendeten Datenbanken (DB) bzw. Datenbank-management Systeme (DBMS) und technische Schnittstellen der im Labor zur Verfügung stehenden ERP-Systeme untersucht.

Ein grundlegender Bestandteil fast aller Informationssysteme stellt die Datenhaltung dar.

Informationssysteme, so auch ERP-Systeme, nutzen je nach Anforderung unterschiedlichste Methoden zur Datenhaltung. Neben dem normalen Dateisystem stellen Datenbanken wohl die am weitesten verbreitete Möglichkeit zur Datenspeicherung dar. Grundsätzlich lassen sich Datenbanken in verschiedene Modelle unterscheiden. Neben klassischen relationalen Datenbanksystemen finden sich zunehmend auch NO-SQL, Key- und Objekt-basierte Datenbanken. Aufgrund ihrer Historie in den 90ern, nutzen ERP-Systeme zu weiten Teilen relationale Datenbanken.

Die im Labor der Universität bereitstehenden ERP-Systeme wurden deswegen auf ihre Datenbanken und Datenspeicher hin untersucht. Bei den Datenbanken von *Microsoft Dynamics NAV, weclapp* und *odoo* handelt es sich um SQL Server. *IFS Applications 10* und *godesys* nutzen primär Oracle DB. Die ERP-Systeme *Canias* und *Xentral* verwenden MySQL. *AvERP* nutzt das freie DBMS Firebird. Außerdem wurden die marktführenden ERP-Systeme Oracle und SAP ausgewertet. Während Oracle weitestgehend auf die hauseigene Oracle DB setzt, bildet SAP ein flexibles Modell, in dem grundsätzlich mehrere Datenbanken als Grundlage betrieben werden können (Oracle, IBM DB2, SAP Hana).

Eine für das Projekt wesentliche Grundlage stellt die Datenhaltung in den Systemen dar. Arbeitspaket C2 umfasste deswegen neben der Analyse der Datenbanken eine Analyse der zugrunde liegenden Datenmodelle.
Hierfür wurden die Datenbankmodelle gängiger ERP-Systeme basierend auf aus den Datenbanken extrahierten ER-Modellen analysiert.
Je nach Größe und Zielgruppe des Systems bestehen die ER-Modelle moderner Systeme aus bis zu mehreren tausend Tabellen, die objektrelational miteinander verknüpft sind. Für die Analyse wurden die Systeme Microsoft Business Central, SAP S/4 Hana und Godesys ERP (bzw. STEPS ERP) fokussiert.

Tabelle 2: Darstellung der Schnittstellen

System	Technische Schnittstelle
Microsoft Dyn. NAV	ODATA-Services / SOAP
Microsoft Bus. Central	ODATA / REST
IFS Applications 10	SOAP (BizAPI)
Canias	SOAP
Xentral	REST-API
AvERP	keine Informationen
godesys ERP	proprietär
STEPS ERP	REST-API
weclapp	REST-API
odoo	XML-RPC / REST-API

Eine Analyse nicht proprietärer Schnittstellen (vgl. Spichale, 2019) erfolgte auf Basis der im ERP-Labor zur Verfügung stehenden Systemlandschaft sowie öffentlichen und von Herstellern eingeholten Informationen. Ein Überblick über APIs vorhandener ERP-Systeme in Tabelle 3 zeigt für ERP-Systeme des ERP-Labors, die im Forschungsprojekt betrachtet wurden, die entsprechend vorhandenen technischen Schnittstellen und die inhärente Heterogenität.

Das wesentliche Ergebnis des ***Teilarbeitspakets C1*** bilden die technischen Rahmenbedingungen, welche die Grundlage für das in AP C5 vorgestellte Testbed bilden.
Die Analyse der Möglichkeiten zur technischen Anbindung einer Add-On-artigen Entwicklung zeigte, dass unterschiedliche Systeme unterschiedliche Schnittstellentechnologien bieten und kein eindeutiger „Standard“ ausgemacht werden kann. Während das Testbed einen unabhängigen Ansatz verfolgen kann, muss für das Add-In eine Designentscheidung zu einem zukunftsfähigen Ansatz getroffen werden. Aufgrund der Häufigkeit und Zukunftsfähigkeit bieten sich hier eindeutig REST-APIs als technologische Schnittstelle an.

3 Teilarbeitspaket C2: Analyse von ERP Systemen

Um Anforderungen für die zu entwickelnde Machine Learning Toolbox bzgl. der Datengrundlage zur Anomalieerkennung zu schaffen, wurden Datenmodelle von ERP-Systemen über mehrere Hersteller hinweg verglichen.
Mit dem Ziel, strukturelle Ähnlichkeiten und Überschneidungen zwischen den Datenmodellen herauszuarbeiten, wurden ER-Modelle unterschiedlicher Datenbanken erstellt und semantisch und syntaktisch analysiert. Untersucht wurden die Datenmodelle aller AP C1 aufgeführten Datenmodelle, wobei ein weitaus größerer Fokus auf den Systemen Godesys (später Step Ahead) SAP S/4 HANA und Microsoft Business Central lag.

Hinsichtlich des Projekts kann festgehalten werden, dass sich ERP-Systeme grundsätzlich gut für die Anomalie-Detektion und Betrugserkennung im betrieblichen Umfeld eignen. Die Speicherung der Daten in meist relationaler Form macht es einfach, große Datenmengen zu extrahieren. Durch die bestehenden Relationen im Datenmodell können Daten in unterschiedlichen Tabellen über deren Fremdschlüssel mit anliegenden Informationen in Verbindung gesetzt werden und so kontextuelle Zusammenhänge beibehalten werden.

Ein Problem für die Generalisierung der Modelle, basierend auf holistischen Ansätzen, stellt die Flexibilität der Systeme in Verbindung mit den genutzten relationalen Datenbanken dar. Softwarehersteller von ERP-Systemen ermöglichen es Anwendern durch Customizing, Funktionen im System an- bzw. abzuschalten und das System so den Unternehmensprozessen entsprechend anzupassen. In Verbindung mit starren Datenbank-Schemata führt diese Flexibilität allerdings zu Datenbankschemata mit sehr vielen ungenutzten Datenbankfeldern und in Folge zu dünn besetzten Datensätzen (sparse data).

Die Analyse der Datenmodelle zeigte, dass sich Modelle unterschiedlicher Systeme teils stark voneinander unterscheiden. Obwohl teils dieselben Daten gespeichert werden, finden

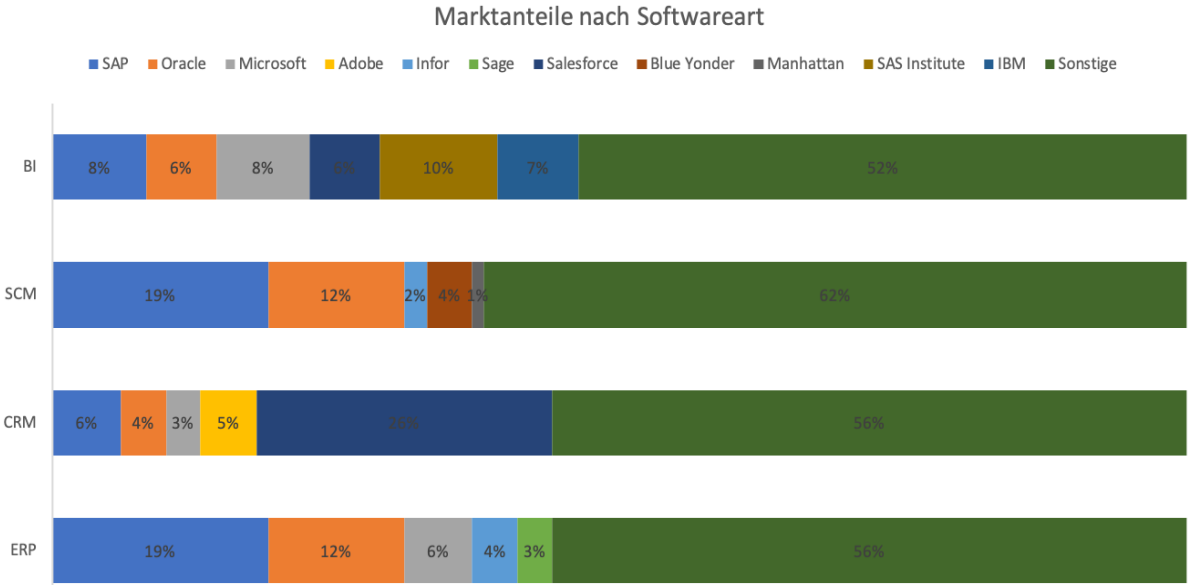

Abbildung 3: Marktanteile

sich für dasselbe Datum als Information in unterschiedlichen Systemen Abweichungen in der Normalisierung und den Datentypen.

Dennoch konnte eine Minimalform ermittelt werden, was die Generalisierung von Machine Learning Modellen, insbesondere hinsichtlich der produktiven Bereitstellung (Deployment), vereinfacht.

4 Teilarbeitspaket C3: Analyse sonstiger Unternehmensdaten

Neben ERP-Systemen existieren im Unternehmensumfeld eine Vielzahl weiterer IS, die möglicherweise durch Verknüpfung von betriebswirtschaftlichen Funktionalitäten als zusätzliche Datenquellen für eine verbesserte Performance führen können oder auf die eine Anwendung zur Anomaliedetektion, im Sinne von Wachstumsfeldern, ausgeweitet werden kann. In C3 wurde untersucht, welche Optionen und Potenziale existieren. Aus einer Literaturrecherche zur Bereichsabdeckung dieser IS Systeme innerhalb eines Unternehmens konnte, wie in Abbildung 1 ersichtlich, eine bereits umfängliche Abdeckung der Unternehmensprozesse durch ERP-Systeme identifiziert werden (vgl. Mertens, 2013; Mertens, Meier, 2009).

Um einen zusätzlichen Überblick über mögliche andere Systeme zu erlangen, wurde eine Marktanalyse für Unternehmenssoftware basierend auf der strategischen Marktanalyse nach

Theobald (2016) durchgeführt. Diese untergliedert sich in die Festlegung des Marktes, die Fixierung der Fragestellung, die Erstellung eines quantitativen und qualitativen Marktprofils, und die Analyse und Bewertung der Informationen (Theobald, 2016).
Als Markt wurde der weltweite Markt für Unternehmenssoftware in der Marktanalyse betrachtet. Im quantitativen Marktprofil wurden die quantitativen Größen wie Umsatz- und Absatz-Volumina ermittelt, um den Fokus der folgenden qualitativen Analyse zu schärfen.

Im Jahr 2020 betrug der Markt für Unternehmenssoftware 210 Milliarden US-Dollar und wird voraussichtlich bis 2026 mit einer CAGR von 8,8% wachsen. Innerhalb des Marktes erzielten die Vereinigten Staaten (105 Mrd. US$) und Deutschland (11 Mrd. US$) den höchsten Umsatz im Jahr 2020.
Das Segment der ERP-Software bildet den größten Teilmarkt für Unternehmenssoftware (Statista, 2020; Färber, 2020; Färber, 2021).
Es gibt eine Vielzahl von Trends auf dem Softwaremarkt, von denen sich die meisten auf bestimmte Softwaretypen beziehen. Neben ERP-Systemen konnten Webshops, SCM und CRM-Systeme als wesentliche Software ausgemacht werden, auf die sich das Ziel der Anomaliedetektion ausweiten lässt (vgl. Abb. 2).

Aufgrund der Vorteile, die durch eine flexible und ortsunabhängige Nutzung sowie durch den Wegfall der lokalen Installation und Wartung entstehen, entwickelt sich Software-as-a-Service (SaaS) immer weiter zum Standard in der Softwarenutzung (Statista, 2021). Laut eines Marktforschungsberichts der Synergy Research Group im Jahr 2020 hat sich das Cloud-Konzept bereits in allen wichtigen Anwendungsbereichen durchgesetzt. Vor allem im CRM-Segment herrscht eine große Nachfrage. Vergleichsweise gering fällt hingegen der SaaS-Anteil bei ERP aus, allerdings steigt dieser kontinuierlich. Die Zunahme von Cyberangriffen auf Cloud-Applikationen (Statista, 2021) bildet dabei ein weiteres Potenzial für den zukünftigen Einsatz und die Relevanz der im Projekt fokussierten Anomalieerkennung.

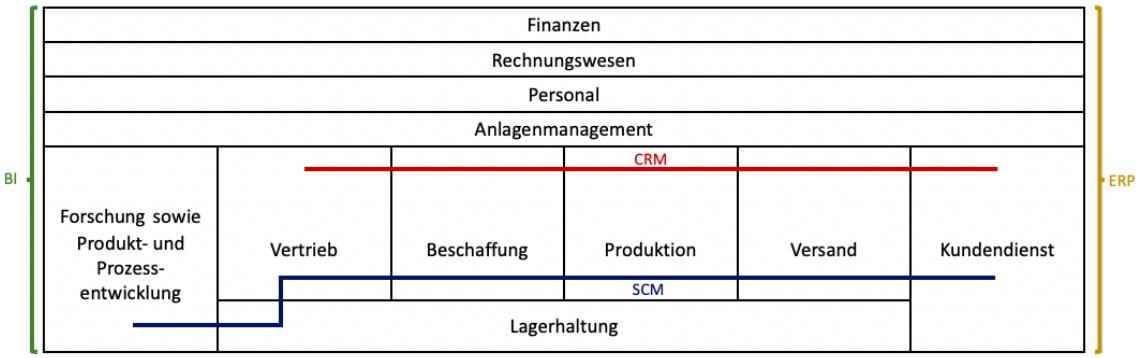

Abbildung 4: Einordnung der Systeme (in Anlehnung an Mertens & Meier, 2009)

Grundsätzlich zeigt die Marktanalyse, dass sich die Betrugserkennung auf weitere betriebliche Daten ausweiten lässt, ERP-Systeme aber den größten Marktanteil hinsichtlich der Verbreitung ausmachen.

5 Teilarbeitspaket C4: Berücksichtigung Datenschutzrechtlicher Anforderungen

Für das Testbed selbst sind keine besonderen datenschutzrechtlichen Anforderungen relevant, da auf unbedenklichen, öffentlichen Daten oder synthetischen Daten gearbeitet wurde.

Die Analyse echter Datensätze wurde von der datenschutz süd im rechtlichen Umfang ebenfalls geprüft, wurde aber auf Basis des §§27 und Art. 89 DSGVO im Forschungsprojekt als unbedenklich eingestuft. Dennoch mussten grundlegende Datenschutzaspekte garantiert werden. Datensätze von Konsortialteilnehmern und Evaluations-partner, die das Projekt im Projektverlauf erhalten haben, wurden deswegen pseudonymisiert, verschlüsselt und lediglich auf Geräten und Servern der Universität verarbeitet (vgl. Arbeitspaket B).

Für den Realbetrieb einer Softwareanwendung zur holistischen Datenanalyse bestehen jedoch rechtlich weitere Anforderungen. Ausnahmeregeln der DSGVO, wie sie im Projekt zur Entwicklung gelten, können nicht in Anspruch genommen werden. Hinzu kommen schärfere gesetzliche Bestimmungen basierend auf DSGVO und weiteren rechtlichen

Rahmenbedingungen aus dem Arbeitsschutzrecht, welche von Unternehmen für den Betrieb jeglicher IT-Systeme geprüft und eingehalten werden müssen[2].
Für den vereinfachten Einsatz und die einfachere Zulassung sieht das Projekt deswegen die Anonymisierung gewisser Daten vor der eigentlichen Analyse vor.

Nach einer Prüfung der geplanten Aktivitäten für synthetische und reale Datensätze und der Gegebenheiten konnten für den Demonstrator keine einschränkenden Rechtsvorschriften ausgemacht werden, welche die Forschungsaktivitäten beeinflussen könnten.
Trotzdem wurden für reale Datensätze von Partnern Schritte zur Anonymisierung festgehalten, die sich aber auf die wesentlichen personenbezogenen Merkmale von Kunden und mitarbeiterbezogene Daten in Datensätzen beschränkten. Die Schritte sind dabei weniger als Resultat aus rechtlichen Vorgaben, sondern vielmehr als eigene Absicherung, und Vertrauensmaßnahmen gegenüber den Forschungs- und Evaluationspartnern zu sehen.

6 Teilarbeitspaket C5: Erstellung eines Testbed

Für Entwicklung einer im Projekt wurde ein Testbed im Sinne einer Data Pipeline nach Stat-of-the-Art Data Engineering Standards implementiert. Abbildung 3 zeigt die Data Pipeline inklusive der verwendeten Open-Source-Software, die zur Extraktion und Speicherung der Daten verwendet wurde. Die Daten wurden entweder über die direkte Anbindung der Systeme per Datenbankexport oder über von den Systemen bereitgestellte technische Schnittstellen (API) angebunden und dann entweder änderungsbasiert oder zeitgesteuert, differentiell über die Pipeline extrahiert.

[2] Ausschlussklausel, Prüfung der jeweils zum Einsatzzeitpunkt gültigen Rahmenbedingungen müssen eigenständig geprüft werden.

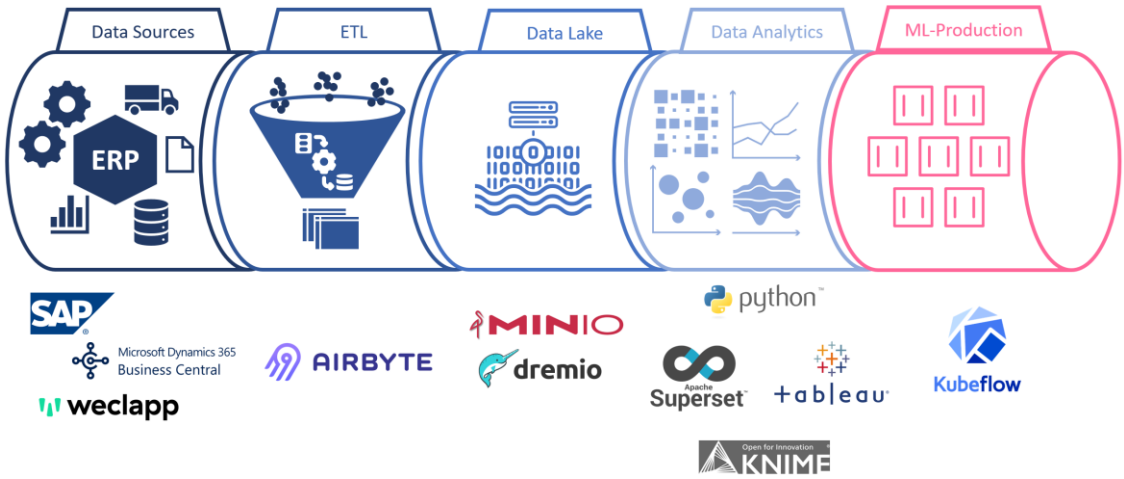

Abbildung 5: Einordnung der Systeme (in Anlehnung an Mertens & Meier, 2009)

Für die Präparation der Daten und die explorative Datenanalyse (EDA) wurden Python-Notebooks sowie Tools wie Apache Superset, KNIME und RapidMiner genutzt. Die Verarbeitung wurde auf den für das Projekt beschafften Servern durchgeführt, welche zur Nutzung in bestehende Infrastruktur integriert wurden. Zur Verarbeitung großer Datenmengen wurde ein kleines Spark Cluster sowie ein HPC mit SLURM bzw. Kubernetes genutzt.

Neben echten Datensätzen von zwei Partnern mussten, insbesondere zur Veröffentlichung und Tests, synthetische Daten generiert werden. Aufgrund der Anforderungen zur Modellierung von möglichst realistischem Betrug wurden unterschiedliche Möglichkeiten zur Datengenerierung getestet.

6.1 Datengenerierung per Serious Game

Eine im Projekt genutzte Möglichkeit zur Datengenerierung ist die Spiel-basierte Generierung von synthetischen ERP-Daten. Dafür wurde im Zuge einer Studie eine spielbare Umgebung entwickelt, ein sogenanntes Serious Game, mit dem realitätsnahe ERP-Daten erzeugt werden können.
Der Prototyp wurde im Sinne eines rundenbasierten kooperativen Multiplayer-Spiels entwickelt. Die Spieler werden durch Geschäftsbeteiligungen auf persönliche Punktekonten

dazu ermutigt, gute Geschäftsentscheidungen zu treffen und das Unternehmen betriebswirtschaftlich gewinnbringend zu steuern.
Gleichzeitig wurde den Spielern die Möglichkeit geschaffen, verschiedene Arten von Betrug zu begehen. Als Spielziel wurde der maximale persönliche monetäre Gewinn festgelegt, ohne sich bei Betrug entdecken zu lassen. Durch die spieltheoretischen Ansätze existiert ein kontinuierlicher Anreiz, gut versteckten Betrug zu begehen. Im Spiel wurde den Spielern eine ERP-typische aber simplifizierte Oberfläche geboten, mit deren Hilfe auch Laien Prozesse des Unternehmens steuern können und welche somit die Datengenerierung für eine breitere Menge an Nutzern ermöglicht. Zusätzlich erlaubt die Oberfläche die Steuerung von Schlüsselvariablen in einem MTS-Produktion-Szenario, um neben der Steuerung der Prozesse des Unternehmens auch möglichst flexibel und kreativ Betrugsfälle zu erzeugen und zu verbergen. Um realitätsnahe Datensätze zu generieren, musste das Balancing mittels realer Parameter angepasst werden.

Das konzeptionierte Spiel ermöglicht insbesondere die Entstehung vieler neuer Betrugsfälle durch die Bereitstellung vieler Freiheitsgrade und menschliche Interaktion.
Als Limitierung muss eine Abwägung zur Menge der Freiheitsgrade und der Zugänglichkeit des Spiels für Laien in Kauf genommen werden. Die Generierung von ERP Daten mittels Serious Game wurde im Rahmen eines Konferenzbeitrags veröffentlicht (Tritscher et al., 2021).
Drittunternehmen haben hierzu bereits Ihr Interesse bekundet (vgl. AP H).

6.2 Datengenerierung im ERP-System

Einen weiteren Ansatz zur Datengenerierung für das Projekt stellte die Datengenerierung mittels menschlicher Probanden dar. Hierbei wurden von Studierendengruppen Datensätze in einem echten ERP-System erzeugt. Insgesamt arbeiteten neun Studierenden Gruppen mit je fünf Nutzern als Probanden in unterschiedlichen Zusammenstellungen an

der Generierung von Datensätzen für das Projekt (vgl. Tritscher 2022).

Datengenerierung im ERP-System:
Für die Datengenerierung wurde das von der Firma HEC MONTREAL entwickelte Unternehmensspiel ERPsim verwendet, das auf SAP S/4 Hana basiert. Die ERPsim Simulationsumgebung stellt damit ein Planspiel bereit, das in einem realen ERP-System stattfindet und ein Unternehmensumfeld über ein Geschäftsjahr hinweg simuliert. Die Studierenden bedienten das ERP-System und bildeten, analog zu einem realen Unternehmen, alle gängigen Prozesse vom Einkauf bis zum Vertrieb nach.

Für das Projekt konzentrierten sich die Probanden insbesondere auf den Purchase-to-Pay Prozess. So konnten mehrere Datensätze mit und ohne Fraud erzeugt werden. Tabelle 3 zeigt die von den Probanden ausgenutzten Fraud-Fälle, welche aus dem ACFE-Jahresbericht (ACFE 2021) abgeleitet wurden.

Tabelle 3: Betrugsfälle in SAP S/4 Hana

Funktion	Szenario	Beschreibung
Einkauf	E1	Änderung des Durchschnitts-Einkaufspreis
	E2	Manuelle Buchung des Wareneingangs und Manipulation
	E3	Veränderung der Lieferdaten
	E4	Veränderung der Kreditorenstammdaten
	E5	Eigenbestellung von Ware
Verkauf	V1	Variation der Verkaufspreise
	V2	Veränderung der Lieferanschrift
Finanzen	F1	Manipulation des Abschreibungsplans für Assets
	F2	Manipulation der Bilanz / Buchungen

Datengenerierung in Godesys:
Zur Datengenerierung in Godesys wurde, anders als bei SAP, auf das komplette System zurückgegriffen. Die folgende Tabelle 3 zeigt die implementierten Fraud-Fälle.

Im Einkaufsprozess konnten die bereits beschriebenen Fraud-Fälle E3 und E4 umgesetzt werden. Im Verkaufsprozess wurden, wie in SAP S/4 Hana, die Fraud-Fälle V1 und V2 implementiert. Im Produktionsprozess konnte die Anzahl produzierter Güter bzw. die nachträgliche Materialentnahme manuell verändert werden (P1).

Tabelle 4: Betrugsfälle in Godesys ERP

Funktion	Szenario	Beschreibung
Einkauf	E1	Veränderung der Lieferdaten
	E2	Veränderung der Kreditorenstammdaten
Verkauf	V1	Variation der Verkaufspreise
	V2	Veränderung der Lieferanschrift
Produktion	P1	Materialentnahme und Veränderung der Produzierten Mengen
	P2	Entnahme über irregulären Ausschuss

6.3 Skriptbasiert synthetische Datengenerierung

Neben den bereits beschriebenen Möglichkeiten zur Datengenerierung wurde im Projekt ebenfalls die skriptbasierte synthetische Datengenerierung genutzt. Dazu wurden für unterschiedliche Belegarten (Bestellung, Eingangsrechnung, etc.) gesonderte Skripte entwickelt und basierend auf Echtdaten Daten aus 6.1 weitere Belege erzeugt, die dem Muster dieser Daten folgten. Innerhalb der Beschaffung wurden so Datensätze für Lieferantenstammdaten, Bestellungen, Lieferantenrechnungen und ausgehende Zahlungen generiert. In der Vertriebsfunktion wurden Datensätze für allgemeine Debitorenstammsätze, Verkaufsaufträge und Auslieferungen generiert.

Stammdaten-Anomalien sowie Tabelleneinträge für die Anomalien, „Abrechnung von nicht erbrachten Leistungen über inaktive Lieferanten“ und „ausgehende Zahlung an eine abweichende Bankverbindung“ mussten aber manuell nachgebildet werden, da diese keinem Muster / keiner Verteilung folgten. Skripte stellten

sich daher als nur bedingt nutzbar für die Datengenerierung im Projekt heraus.

6.4 Deep Learning basierte synthetische Datengenerierung

Zur Datengenerierung wurden im Projekt außerdem Methoden der Künstlichen Intelligenz evaluiert. Sogenannte Generative Adversarial Networks (Goodfellow et al. 2014) werden bereits zur Erzeugung realistischer Bilder erfolgreich genutzt, können aber auch bei der Erzeugung von relationalen Daten eingesetzt werden (Xu et al.2018 und Xu et al. 2019). Im Projekt wurden einige aktuelle Arbeiten zur Nutzung von GANs für die synthetische Generierung von Netzwerkdaten (vgl. Ring et al. 2019) sowie von tabellarischen Daten implementiert und mit Variational Auto-Encodern (VAE, vgl. Kingma & Welling 2013) verglichen, hinsichtlich ihrer Fähigkeiten Finanztransaktionsdatensätze variierender Komplexität nachzubilden.

Bei der Analyse der Modelle zeigte sich, dass sowohl VAE als auch GAN basierte Modelle die Verteilungen und Abhängigkeiten der Datensätze effektiv erfassen und in der Lage sind, einen Datensatz zu synthetisieren, der den Merkmalen des zugrunde liegenden realen Datensatzes entspricht auf dem trainiert wurde.

Im Vergleich neigt das GAN-Modell dazu, numerische Verteilungen genauer zu erfassen, während das VAE-Modell weniger anfällig für fehlende Merkmalswerte ist, was typischerweise als "Mode-Kollaps" (vgl. Mescheder 2018) bezeichnet wird, oft als genereller Nachteil von GAN-basierten Modellen beschrieben.

7 Zusammenfassung der Ergebnisse des Arbeitspakets C

Insgesamt konnten mit dem AP C wesentliche Grundlagen für das Projekt erarbeitet werden.

In Arbeitspaket C1 und C2 wurden dazu die technischen Grundlagen für das Testbed erarbeitet, aber auch erste Anforderungen für die zu entwickelnde Toolbox und den Demonstrator herausgearbeitet.

Obwohl in der Analyse sonstiger Unternehmensdaten festgestellt wurde, dass eine Betrugserkennung auf Basis

von ERP-Systemdaten die wesentlichen und weitverbreitetsten Prozesse im Branchendurchschnitt abdeckt und demnach eine Hinzunahme weiterer Systeme wohl zu keinen wesentlichen Verbesserungen der Ergebnisse führen würde, konnten durch die Analyse der Zukunftstrends im Softwaremarkt, mögliche Anschlussmöglichkeiten für das Projekt ausgemacht werden.

Basierend auf den erarbeiteten technischen Grundlagen aus AP C1 und C2, konnte in AP C5 ein Testbed analog einer Datenpipeline entwickelt werden, um die in AP D und AP E betrachteten Verfahren zu entwickeln und evaluieren.
Basierend auf unterschiedlichsten Verfahren zur Datengenerierung wurden außerdem diverse synthetische Datensätze unterschiedlicher Qualität erzeugt. Neben Skripten und Simulationen kamen hier auch neuronale Netze in Form von GANs zum Einsatz.

8 Literaturverzeichnis[3]

- *ACFE (2021) Association of Certified Fraud Examiners. "2021 Report To Members." 25 Aug. 2022, www.acfe.com/fraud-resources/report-to-members.legacy.acfe.com/report-to-the-nations/2022/?_ga=2.193760173.1024481359.1661412185-661814964.1661412183.*
- *Färber, J. M. (2020): Enterprise Software Report 2020. Statista. https://de.statista.com/statistik/studie/id/85016/dokument/enterprise-software-report/.*
- *Färber, J. M. (2021): Software Report 2021. Statista. https://de.statista.com/statistik/studie/id/102693/dokument/software-report/.*
- *Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems 27 (2014).*

[3] Eigene Publikationen wurden fettgedruckt gekennzeichnet

- *Mertens, P. (2013): Integrierte Informationsverarbeitung. Operative Systeme in der Industrie. 18. Aufl., Springer, Wiesbaden.*
- *Mertens, P.; Meier, M. (2009): Integrierte Informationsverarbeitung. Planungs- und Kontrollsysteme in der Industrie 2. 10. Aufl., Springer, Wiesbaden.*
- *Mescheder, Lars Geiger, Andreas Nowozin, Sebastian: Which Training Methods for GANs do actually Converge? 13. Januar 2018, arxiv:1801.04406.*
- ***Ring, Markus, et al.** "Flow-based network traffic generation using Generative Adversarial Networks." Computers & Security, vol. 82, 1 May. 2019, pp. 156-72, doi:10.1016/j.cose.2018.12.012.*
- *Spichale, Kai. API-Design, 2nd Edition. Dpunkt Verlag, 2019.*
- ***Tritscher, Julian, et al.** "2021 IEEE Conference on Games (CoG)." A financial game with opportunities for fraud. IEEE, 17 Aug. 2021, pp. 1-5, doi:10.1109/CoG52621.2021.9619070.*
- ***Tritscher, Julian, et al.** "Open ERP System Data for Occupational Fraud Detection." arXiv preprint arXiv:2206.04460 (2022).*
- *Xu, Lei, and Kalyan Veeramachaneni. "Synthesizing tabular data using generative adversarial networks." arXiv preprint arXiv:1811.11264 (2018).*
- *Xu, Lei, et al. "Modeling tabular data using conditional gan." Advances in Neural Information Processing Systems 32 (2019).*

Arbeitspaket D: Analyse und Modellierung von Vorgängen in ERP-Systemen

Überblick:

Arbeitspaket D befasst sich im Wesentlichen mit der Datengrundlage auf Basis der im Projekt fokussierten Ansätzen zur Betrugserkennung. Neben der Analyse der Eignung der Daten und Aufbereitung der Datengrundlage werden erste Ergebnisse bzgl. Der Embeddingmethoden zusammengefasst und Ergebnisse bzgl. der Generalisierbarkeit der Verfahren diskutiert.

Inhaltsverzeichnis Arbeitspaket D

1 Methodik, Vorgehensweise und Unterteilung in Teilarbeitspakete

Arbeitspaket D befasst sich mit in Informationssystemen (IS) abgebildeten Vorgängen. Für die Anomalie Detektion spielen Vorgänge in Informationssystemen, wie ERP-Systemen, eine wesentliche Rolle, da es sich dabei um die semantische und kontextualisierte Abbildung aller von Nutzern ausgeführten Aktivitäten handelt. Einer Repräsentation der Vorgänge als Daten wird demnach ein großes Potenzial zur Nutzung in der Anomalieerkennung zugeschrieben.

Teilarbeitspaket D1, die Anforderungsanalyse, bildet die Grundlage, durch die die Anforderungen an die Datenerhebung der Vorgangsdaten strukturiert erhoben und formal dokumentiert werden. Hierzu werden in Arbeitspaket D1 inhaltliche und qualitative Anforderungen aufgestellt, basierend auf den im Fokus stehenden Untersuchungsmethoden und Modellen.

Teilarbeitspaket D2 untersucht, wie Repräsentationen (bspw. von betrieblichen prozessualen Abläufen) in Machine Learning Modelle eingebettet werden können. Hierzu wurden Repräsentationen mittels unterschiedlicher Ansätze in gängige ML Modelle integriert und über Experimente die Effizienz der Methoden für die Betrugserkennung untersucht.

Teilarbeitspaket D3, beschäftigt sich mit der Anwendung von hypothesengetriebenen Ansätzen zur Erklärung und Modellierung von ERP-Vorgängen mit Hilfe von Expertenwissen.

Teilarbeitspaket D4, die Substrukturanalyse und Modellierung von ERP-Systemen, wurde im Laufe des Projekts unterteilt. Dabei wurde einerseits die Methodik von Subtrails und andererseits ein auf Process Mining aufbauender Ansatz, der Methoden des Machine Learning inkludiert, untersucht.

2 Teilarbeitspaket D1: Anforderungsanalyse

Teilarbeitspaket D1 beleuchtet die in IS wie ERP-Systemen vorliegenden Daten unter Berücksichtigung der im Projekt fokussierten Ansätze zur Anomalie-Erkennung (Anomaly Detection; AD). Darauf aufbauend erfolgte die systematische strukturierte Ableitung von Anforderungen.

Um eine spätere Auswertung für unterschiedliche methodische Ansätze zu schaffen, wurde in Teilarbeitspaket D1 die Datengrundlage untersucht, wobei der Fokus getrennt auf die Datenmodelle der Systeme und vorliegende Prozessdaten gerichtet wurde.

Im Projekt wurden einerseits Datenbank Schemata, aber auch ableitbare Prozessdaten und weitere exportierbare Formate analysiert.

Für die Anwendung von Machine Learning Verfahren wurden im Rahmen des Projekts Rohdaten transaktionaler Art, also Bewegungsdaten, welche bei alltäglichen Geschäftsvorgängen wie Bestellungen und Rechnungen anfallen, als geeignet bewertet. Für Hyp- / Subtrails hingegen wurde die Analyse von ableitbaren Prozessrepräsentationen genutzt auf Aktivitätsebene besonders relevant, da nur diese den Geschäftskontext beleuchten und damit die nötigen betriebswirtschaftlichen Daten umfassen.

Für beide Bereiche wurden Daten über unterschiedliche Systeme hinweg analysiert.

Obwohl die Analyse der Datenbankschemata ein sehr heterogenes Bild zeichnete, konnten Überschneidungen in den Datenmodellen ausgemacht werden. Wesentlich größeres Problem stellen allerdings die in unterschiedlicher Form und divers benannten Daten dar. Obwohl die Systeme aufgrund des betriebswirtschaftlich fast einheitlichen Kontexts mit denselben Daten arbeiten, werden diese unterschiedlich benannt und liegen oft auch mit unterschiedlichen Datentypen und Datenbank Normalisierungsformen vor.

Da manuelle Modellierung aufwändig und nur unter Umständen zu validen Modellen führt, wurde für die Anwendung von Hyp-/ SubTrails die datenbasierte Extraktion von

Prozessmodellen aus IS untersucht. Hierzu wurden insbesondere Process Discovery Verfahren aus dem Process Mining Umfeld auf ihre Anwendbarkeit untersucht.

Zusätzlich wurden in ersten Versuchen Exportformate wie Übertragungsstandards (EDIFACT) und Audit-Report analysiert, da deren, per Definition einheitliche Struktur, eine Nutzung hätte vereinfachen können.

Bei den Untersuchungen des Datensatzes mit EDIFACT-Daten hat sich gezeigt, dass die Daten zwar einheitlich beschrieben sind, die reale Nutzung aber stark variiert. Bei der Analyse der Audit-Reports hingegen hat sich gezeigt, dass die Daten zu stark aggregiert und vereinfacht werden. Sowohl EDIFACT Daten als auch Audit-Reports konnten deswegen im Projekt nicht weiterverwendet werden.

In der folgenden Tabelle sind die in diesem Teilarbeitspaket erarbeiteten Anforderungen in Bezug auf die zwei wesentlichen Modellierungsmethoden zusammengefasst.

Tabelle 1: Identifizierte Anforderungen

	Anforderungen zur ML basierten Modellierung
1	**Exportierbarkeit** der Daten, in der aktuellen Ausprägung in Flat Files oder komplexe Datenformate in Parquet um bspw. Datentypen zu erhalten.
2	**Entitätsbasiertes / relationales Datenmodell**, um wesentliche Daten über einen Geschäftsvorfall zusammenführen zu können
3	Flache Datenstruktur / Auflösbarkeit, um komplexe 1:n und n:n Beziehungen auflösen zu können.
4	Möglichkeit zur **Aggregation** der Daten
5	**Datenmenge**, insbesondere hinsichtlich saisonaler Schwankungen und zeitlichen Veränderungen in Unternehmen sollten Datensätze einen größeren Zeitraum umfassen, in dem Muster wiederholt vorkommen.
	Anforderungen zur hypothesenbasierten Modellierung
5	Exportierbarkeit der Daten, in der aktuellen Ausprägung in Flat Files oder komplexe Datenformate in Parquet, um bspw. Datentypen zu erhalten.
6	Fähigkeit zur Aggregation der Daten in Graphen

7	Datenabhängigkeiten, Restrukturierung der Graphen, sodass Modellierung als Markov Kette 1. Ordnung notwendige Abhängigkeiten erhält.
8	Formulierbarkeit von Hypothesen für Graphen Transition (Zustandswechsel)
9	Formulierung über Kantenverteilungen muss aus praktischen Überlegungen und Expertenwissen abgeleitet werden können
10	Bei der Nutzung von Process Mining zur Generierung der Graphen für hypothesenbasierte Ansätze muss eine durchgängige Case ID, und ein Zeitstempel zuordenbar sein

3 Teilarbeitspaket D2: Semantische Einbettung von Abläufen in ERP-Systeme

Die semantische Einbettung bezeichnet die Integration von Kontextinformationen in Ansätze zur Anomalieerkennung. Im Fokus des Projekts wurde sich auf Methoden konzentriert, die eine Einbettung der prozessualen Abläufe ermöglichen. Zur semantischen Einbettung wurden Daten aus ERP-Systemen des Testbeds mittels Representation Learning in Modelle integriert und verschiedene Vorgänge innerhalb der Daten erfolgreich gruppiert und klassifiziert.

Dazu wurden etablierte Methoden des Representation Learning (Word2Vec, FastText, GloVe, Paragraph2Vec, Meta Modell) für die Modellierung von ERP-Prozessdaten adaptiert und auf die Daten angewandt. Basierend auf den Einbettungen wurden Autoencoder Architekturen trainiert, mit deren Hilfe die Repräsentationen der Daten evaluiert wurden.

Für alle Methoden wurden unterschiedliche Konfigurationen auf ihre Eignung zur semantischen Einbettung sowohl intrinsisch als auch extrinsisch evaluiert:

Bei der intrinsischen Evaluation haben sich Word2vec basierende Verfahren hervorgetan, semantisch sinnvolle Einbettungen zu liefern (z.B. hinsichtlich Transaktionscode).

Die extrinsische Evaluation befindet GloVe basierte Einbettungen als am besten geeignet für AD mittels Machine- und Deep Learning.

Um Zusatzwissen durch Ontologien aus Datenbank-Schemata und

Workflows der ERP-Systeme zu inkludieren, kam die Modellierung komplexer Zusammenhänge durch Process Mining zum Einsatz. Hierbei kam formalisiertes Zusatzwissen in Form von SQL-Datenbank Schemata zum Einsatz, um von der Transaktionsebene zu abstrakten Darstellungen auf Prozessebene zu gelangen.

Während für ML und AD Verfahren die Modellierung durch Embeddings am besten geeignet war, erforderten hypothesenbasierte Methoden eine abstrahierte Repräsentation beispielsweise durch semantisch bedeutungsvolle Prozessgraphen, damit Hypothesen mit Expertenwissen formuliert werden können.

4 Teilarbeitspaket D3: Entwicklung von Hypothesen zur Erklärung & Modellierung

Basierend auf Erkenntnissen von D1 und D2 sind für die Hypothesenentwicklung abstrakte Repräsentation erforderlich, welche die Daten ausreichend genau erfassen, aber trotzdem Hypothesenformulierung unabhängig von einzelnen Transaktionen ermöglichen. Das ist insbesondere unter Berücksichtigung von Anforderungen aus D1 (vgl. Tabelle 1. vgl. Graph-Repräsentation mit Markov-Eigenschaft 1. Ordnung) schwierig, da entweder Abhängigkeiten zwischen Prozessschritten trivial werden oder Zusammenhänge über Schritte hinweg verloren gehen.

Für die Evaluation der Modelle wurden verschiedene Business Szenarien in unterschiedlichen Detaillierungsgraden modelliert und entsprechende Hypothesen dazu aufgestellt, bspw. über die Entwicklung von Verkaufszahlen, zu Rabattaktionen oder in Hinblick auf die Preisgestaltung.
Durch die Modellierung als De Bruijn Graphen (vgl. De Bruijn, 1946) konnten auch Markov'sche Abhängigkeiten über die erste Ordnung hinaus berücksichtigt werden.
Insgesamt hat sich in mehreren Experimenten gezeigt, dass die Modellierung von Businessszenarien mit HypTrails auf einer groben Abstraktionsebene möglich ist. Auf feingranularen Abstraktionsebenen konnte dagegen, aufgrund des linearen Prozessablaufs und der Vielzahl an beteiligten

Attributen, keine ausreichend präzise Modellierung erreicht werden.

5 Teilarbeitspaket D4: Substrukturanalyse und Modellierung von ERP-Systemen

Zur Substrukturanalyse wurde als Vorstufe für die Anwendung von SubTrails eine Prozessmodellierung und Analyse auf Basis von Process Mining Ansätzen genutzt. Da sich auch aus dem Process Mining Ergebnisse ableiten lassen, werden die Ergebnisse im Folgenden getrennt beschrieben.

5.1 Substrukturanalyse mittels Process Mining und Machine Learning

In Untersuchungen zeigt sich, dass sich einige Betrugsfälle prozessabhängig über mehrere Aktivitäten / Transaktionen erstrecken. Process Mining dient in der Regel dazu, Geschäftsprozesse eines Unternehmens zu analysieren. In zahlreichen Experimenten konnte ebenfalls eine gute Eignung im Bereich Betrugserkennung festgestellt werden. So existiert bereits diverse Forschung über den Einsatz von Process Mining Insbesondere im Umfeld der Wirtschaftsprüfung, der Daten Forensik und dem Auditing (vgl. Krcmar & Baader, 2018; Hosseinpur & Jans 2018; Zerbino et al., 2018)

Im Projekt wurden Modelle des Unsupervised Machine Learnings zur Weiterentwicklung der gängigen Process Mining Ansätze zur AD betrachtet. Die aus Prozess Mining entstandenen abstrakten Prozessschritte sollen hierdurch hinsichtlich ihrer Substruktur auf Transaktionsebene verfeinert werden. Unter Rückbezug auf das systemimmanente Datenmodell und unter der Anwendung von Clustering-Verfahren auf Transaktionsfeatures konnten Substrukturen herausgearbeitet werden.

Die Idee hinter dem Ansatz ist, dass sich ähnliche Transaktionen in kontextspezifische Cluster gruppieren lassen und dementsprechend einheitliche Substrukturen auf Prozessebene abbilden (Gwinner et al., 2023).

5.2 Substrukturanalyse mittels Sub-Trails

Als zweiten Ansatz wurde eine Substrukturanalyse auf Basis von hypothesenbasierten Ansätzen mit Sub-Trails untersucht.

Hierfür wurden aufbauend auf der Analyse mit Process Mining, abstrakte Repräsentationen des Prozessgraphen verwendet, um einzelne Prozessschritte als Pfad zu repräsentieren. Attribute, die einzelnen Transaktionen zugeordnet werden können und nach Expertenwissen hinsichtlich möglicher Substrukturen interessant sind, wurden aus den Rohdaten zusammengeführt und die abstrakte Prozessebene um diese ergänzt. Die so geschaffene Datenbasis wurde schließlich anwendungsspezifisch angepasst und eine Substrukturanalyse basierend auf Subtrails durchgeführt.

In einer Evaluation durch Experten wurden einige der dabei aufgetanen Subgruppen als interessant bewertet, primär jedoch in Hinblick auf unterschiedliche Produktions- und Verkaufsstrategien und weniger für die Erkennung sicherheitskritischer oder anderer ungewöhnlicher Ereignisse.

6 Teilarbeitspaket D5: Adaptionsfähigkeit und Generalisierbarkeit

Hinsichtlich der in D2 betrachteten semantischen Einbettungen wurden im Projekt zwei mögliche Ansätze zur Bewertung der Adaptionsfähigkeit und Generalisierbarkeit analysiert.

Einerseits werden, durch die hohe Abstraktionsgüte auf Prozessebene, eine Datensatz und ERP-System übergreifende Perspektive eingenommen. D.h. für die Anpassung, auf bspw. andere ERP-Systeme, muss lediglich die Modellierung der abstrakten Repräsentationen ausgetauscht werden.

Das wird durch Anpassen oder Austauschen der SQL-Datenbankschemata zur Erzeugung der Repräsentationen erreicht, oder durch geeignete Abbildung der gelernten Embedding Repräsentationen bzw. Transaktionen.

Andererseits kann die Generalisierbarkeit auch durch Beschränkungen auf gemeinsame Attribute erreicht werden. Hierzu wurden Datenmodelle mehrerer ERP-Systeme auf Überschneidungen untersucht. Hierbei zeigten sich zwar Features, die fast jedes System nutzt, unter

Berücksichtigung der unterschiedlich vorliegenden Datentypen und Normalisierung in verschiedenen ERP-Systemen existieren allerdings auch diverse eminente Unterschiede. Hinzu kommt der unterschiedlich gestaltete Umgang mit anwenderspezifischen Einstellmöglichkeiten (Customizing).

Außerdem zeigte sich, unter Berücksichtigung der Ergebnisse aus den Machine Learning Experimenten, dass Anomalien oft durch Auffälligkeiten in vom Customizing abhängigen und damit systemspezifischen Features ausgemacht werden können, was zusätzlich gegen eine systemübergreifende Generalisierbarkeit der Modelle durch eine Einschränkung auf gemeinsame Attribute spricht.

Grundsätzlich lässt sich also ein generisches Datenmodell schaffen, das ein AD-Modell für möglichst viele ERP-Systeme ermöglicht. Allerdings zeigen unsere Ergebnisse, dass die zu erwartende Präzision der Modelle dadurch wesentlich schlechter ausfällt.

Die Generalisierbarkeit von Hypothesen auf Prozessgraph Ebene hingegen wurde positiv bewertet, da sich angepasste Process Mining Modelle aus den Systemen durch bestehende Process Mining Methoden ableiten lassen, wodurch ein generalisierter, abstrakter und bedeutungsvoller Prozessgraphen für Prozesse konstruiert werden kann.

7 Zusammenfassung der Ergebnisse

Im Rahmen des Arbeitspakets wurden in Teilarbeitspaket D1 strukturierte Anforderungsanalysen für die Modellierung von Vorgängen und Prozessen in ERP-Systemen erstellt und in Teilarbeitspaket D2 erfolgreich Repräsentationen von ERP-System Daten auf unterschiedlichen Abstraktionsebenen für die Manipulationserkennung mittels Verfahren des maschinellen Lernens und Hypothesenbasierten Methoden erstellt.

In Teilarbeitspaket D3 wurden Businessszenarien auf hoher Abstraktionsebene erfolgreich hypothesenbasiert modelliert, wobei Schwierigkeiten bei feingranularer Modellierung der Prozesse zu verzeichnen waren.

Teilarbeitspaket D4 beinhaltete die dynamische Gewinnung von

Substrukturen durch Anwendung von SubTrails, wobei wirtschaftlich interessante Subgruppen identifiziert werden konnten, die allerdings weniger auf die Erkennung ungewöhnlic

her Ereignisse abzielen.

Die in Teilarbeitspaket D5 durchgeführte Untersuchung der Adaptionsfähigkeit und Generalität der Ansätze über ERP-Systeme hinweg zeigte eine gegebenenfalls geminderte Performanz bei der direkten Übertragung aufgrund von Customizing und teils fehlenden Attributen. Für die Übertragbarkeit von SubTrails-basierten Ansätzen wurden allerdings Prozessgraphen des Process Mining als gemeinsame Schnittstelle unterschiedlicher ERP-Systeme identifiziert.

8 Literaturverzeichnis[4]

- *De Bruijn, Nicolaas Govert. "A combinatorial problem." Proc. Koninklijke Nederlandse Academie van Wetenschappen. Vol. 49. 1946.*
- *Gronau, N. (2014): Enterprise Resource Planning: Architektur, Funktionen und Management von ERP-Systemen. 3. Aufl., De Gruyter Oldenbourg, München.*
- ***Gwinner, F., Schaschek, M., (2023)** A Machine Learning based Approach for Process Mining as Anomaly Detection in Homogeneous Process Environments, [Forthcomming].*
- *Koschmider, A., Mannhardt, F. and Heuser, T.. "On the contextualization of event-activity mappings." International Conference on Business Process Management. Springer, Cham, 2018.*
- *Loshin, D. (2008): Master Data Management. In: Elsevier Science & Technology Books, Burlington.*
- *Schemm, J. (2008): Stammdatenmanagement zwischen Handel und Konsumgüterindustrie – Referenzarchitektur für die überbetriebliche Datensynchronisation. Dissertation, Universität St. Gallen, Difo-Druck, Bamberg.*

[4] Eigene Publikationen wurden fettgedruckt gekennzeichnet

Arbeitspaket E: Nachvollziehbare Erkennung von Manipulationen

Überblick:

Arbeitspaket E befasste sich mit der Entwicklung von Ansätzen des maschinellen Lernens zur nachvollziehbaren Erkennung von Manipulationen in ERP-System Daten. Hierfür wurden Distanzmaße und Metric Learning Ansätze, hypothesenbasierte Ansätze, sowie End-to-end Ansätze basierend auf neuronalen Netzen auf ihre Eignung zur Erkennung von Manipulationen untersucht und geeignete Ansätze weiterentwickelt.

1 Methodik, Vorgehensweise und Unterteilung in Teilarbeitspakete

Ziel dieses Arbeitspakets E ist es, eine fundierte Grundlage für die folgenden Arbeitspakete und die Anwendung der Verfahren in der Machine-Learning Toolbox zu schaffen. Dabei sollen zum einen die Rahmenbedingungen und Voraussetzungen erarbeitet werden, zum anderen verschiedene Verfahren des maschinellen Lernens zur Anwendung gebracht werden, um anomales Verhalten in ERP-Systemen zu entdecken. Weiterhin soll die Adaptionsfähigkeit und Erklärbarkeit der Modelle untersucht werden. Für die Untersuchungen stehen Rechenressourcen des Projekts durch beschaffte Serverarchitekturen zur Verfügung, die zur Entwicklung und Evaluation der Betrugserkennungsverfahren genutzt werden. Arbeitspaket E unterteilt sich dazu in die Teilarbeitspakete E1 – E5.

Teilarbeitspaket E1 umfasste eine Anforderungsanalyse. Dazu wurden, mittels Qualitativer Interviews von Fachexperten aus den Bereichen Auditing, Wirtschaftsprüfung und Buchhaltung, die Anforderungen für die im Projekt vorgesehene Softwarekomponente zur Betrugserkennung erarbeitet und daraus spezifische Anforderungen auf algorithmischer Ebene abgeleitet. Zusätzlich wurden Ausgangssituation, aktueller Prozess einer Wirtschaftsprüfung und verwendete Softwaresysteme zur Betrugserkennung aufgenommen. Die Ergebnisse aus den qualitativen Interviews wurden mit einer durchgeführten Literaturanalyse zusammengeführt, um schlussendlich eine praxisnahe und dem wissenschaftlichen Stand entsprechende Grundlage für das Projekt zu schaffen.

In Teilarbeitspaket E2 stehen Distanzmaße und Metric Learning im Vordergrund. Distanzmaße sind wichtiger Bestandteil von verschiedenen distanzbasierten Cluster- und Klassifikationsverfahren und können durch geeignete Wahl oder Adaption durch Metric Learning so das Ergebnis verbessern. Hierfür wurden in diesem Teilarbeitspaket mehrere Experimente mit verschiedenen Daten- repräsentationen, Distanzmaßen und

Metric Learning Methoden durchgeführt.

Hypothesenbasierte Ansätze zur Manipulationserkennung wurden in E3 evaluiert. Hierfür wurde HypTrails und Subtrails genutzt und verschiedene Modellierungsperspektiven (analog zu D3) erprobt. Es wurden empirische Untersuchungen zu diesen Modellierungsansätzen durchgeführt sowie theoretische Überlegungen angestellt, die zusammen mit Erkenntnissen aus der Literatur eine nur bedingte Eignung hypothesenbasierter Ansätze in der Praxis nahelegen.

Teilarbeitspaket E4 behandelte die Untersuchung von Generalisierbarkeit. Um Daten aus spezifischen Systemen und darin enthaltene Einflussgrößen gezielt beeinflussen zu können, wurde eine agentenbasierte Simulation konstruiert und darauf aufbauend verschiedene Fraud-Szenarien in unterschiedlichen Business Settings simuliert. In mehreren Experimenten wurden die hypothesenbasierten Ansätze aus E3 in diesem synthetischen Setting auf ihre Generalisierbarkeit über Datensätze und damit Systeme hinweg untersucht.

Im Vergleich zu den hypothesenbasierten Ansätzen werden in Teilarbeitspaket E5 Black-Box Varianten in Form von neuronalen Netzen evaluiert.
Hierfür wurde im Rahmen dieses Projekts eine spezielle Neuronale Architektur (iNALU) weiterentwickelt, die arithmetischen Beziehungen präzise erfassen kann und in mehreren Experimenten zur Betrugserkennung in Finanztransaktionsdaten evaluiert. In weiterführenden Experimenten wurde die iNALU Architektur in andere Black-Box Ansätze integriert und für die Anomalieentdeckung auf ERP Daten ausgewertet. Um die Erklärbarkeit wiederherzustellen, wurden

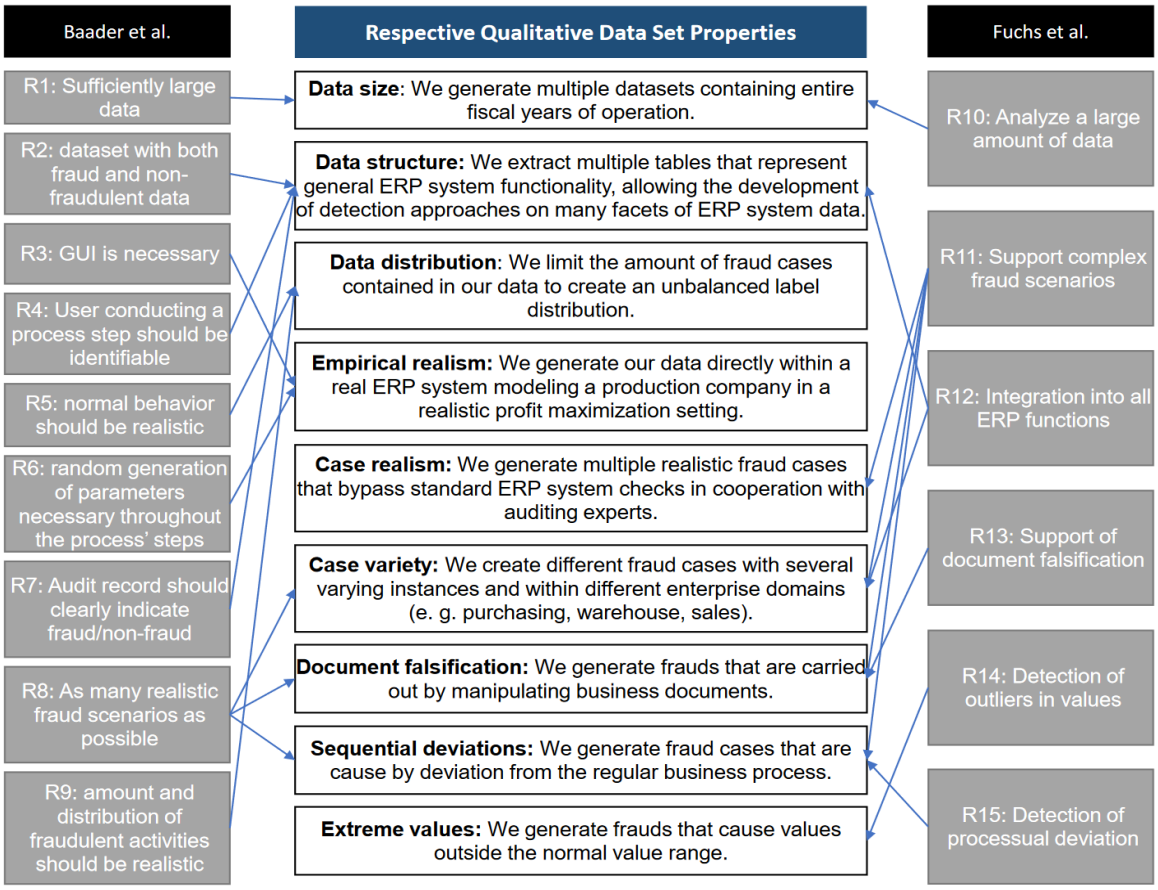

Abbildung 6: Anforderungen und abgeleitete Anforderungen für die Datengrundlage (Tritscher 2022b)

mehrere Post-Hoc Methoden der erklärbaren künstlichen Intelligenz implementiert, adaptiert und in mehreren Experimenten auf ERP Daten bewertet.

2 Teilarbeitspaket E1: Anforderungsanalyse

Um das Verhalten der Algorithmen zur Betrugserkennung nachvollziehbar gestalten zu können, wurden in E1 die diesbezüglichen Anforderungen abgeleitet, die berücksichtigende Maßnahmen erfasst und dokumentiert. Dazu wurden im ersten Schritt die Requirements von **Fuchs et al. (2021)** und Baader et al. (2018) und Cirqueira et al. (2021) kombiniert und durch qualitative Interviews mit den Projektpartnern evaluiert und detailliert. Basierend auf den Anforderungen für Betrugserkennungssysteme wurden außerdem entsprechende Anforderungen für die Daten der in E2, E3, E4 und E5 entwickelten Verfahren abgeleitet (vgl. Abb. 1).

3 Teilarbeitspaket E2: Distanzmaß und Metrik Learning zur Manipulationserkennung

Zunächst wurden in diesem Teilarbeitspaket unterschiedliche etablierte Distanzmaße, sowie Metrik Learning Ansätze aus der Literatur für die distanzbasierte Erkennung von Manipulationen gesammelt.

Um die unterschiedlichen Distanzmaße und Metrik Learning Ansätze auf ihre Eignung zu untersuchen, wurde deren Einsatz in einer Betrugserkennung mittels des k-nächste-Nachbarn Algorithmus untersucht. Die Ansätze wurden sowohl auf synthetischen Rohdaten als auch auf den in Arbeitspaket D2 entstandenen Repräsentationen angewendet und evaluiert. Insgesamt wurden verschiedene etablierte Distanzmaße, sowie unterschiedliche Metrik Learning Verfahren (De Vazelhes, 2020) auf ihre Betrugserkennungsleistung mit verschiedenen in Teilarbeitspaket D2 entstandenen Datenrepräsentationen untersucht.
Insbesondere die durch Autoencoder Netzwerke gelernten Repräsentationen konnten hierbei mit Distanzmaßen wie der Maximumsmetrik gute Ergebnisse bei der Erkennung von Manipulationen erzielen.

Obwohl unterschiedliche Metrik Learning Ansätze auf verschiedenen Einbettungsrepräsentationen angewendet wurden, konnten die besten Resultate der Distanzmaße dadurch nicht weiter übertroffen werden. Insgesamt konnten mittels Distanzmaßen und Metrik Learning einige der Betrugsfälle verlässlich erkannt werden. Eine zufriedenstellende Performanz auf allen Betrugsfällen wurde allerdings nicht erreicht.

4 Teilarbeitspaket E3: Hypothesenbasierte Ansätze zur Manipulationserkennung

Analog zu AP D3 wurden in diesem Arbeitspaket einige Business Szenarien mit unterschiedlichem Detailgrad als Prozessgraphen modelliert.

Für die abstrakte Modellierung, wie sie für D3 gut funktioniert hat ergab sich jedoch, dass die Formulierung

von Fraud-Fällen auf dieser groben Ebene nicht möglich ist, da entsprechende Transitionen im Prozessgraph für diesen Detailgrad nicht von legitimen Transaktionen abgegrenzt werden konnten, ohne als solche trivial zu erkennen zu sein (beispielsweise durch Fraud spezifische Subpfade, die in einem realistischen Setting nicht bekannt sind). Eine feingranulare Modellierung erlaubt dagegen die Betrugsfälle im Prozessmodell abzubilden, bringt jedoch die bereits in D3 diskutierten Probleme durch Markovsche Abhängigkeiten über die 1. Ordnung hinaus mit sich.
Eine Modellierung als De Bruijn Graph konnte hier leider kein zufriedenstellendes Ergebnis liefern, insbesondere auch weil durch die Seltenheit von Betrugsfällen für formulierte Fraud-Hypothesen nicht genügend Evidenz erzeugt werden kann, unabhängig ob solche Fälle in den Daten vorkommen oder nicht.

Nach Bartkowiak 2011 (aber auch nach Amarbayasgalan 2018; Domingues 2019; Oosterlinck 2020) charakterisieren sich Fraud Fälle neben ihrer Seltenheit in der Regel durch ihre Neuartigkeit und die Absicht dahinter diese zu verschleiern. Aus diesem Grund und den aus diesem Arbeitspaket abgeleiteten empirischen Resultaten, erscheint eine Modellierung der Fraud Fälle als Hypothesen als in der Praxis weniger geeignete Lösung da die Neuartigkeit solcher Fälle jeweils auch durch fallspezifisch neu formulierte Hypothesen abgebildet werden müssten.

Nach der in diesem AP vorgeschlagenen detaillierten Modellierung können wohldefinierte Fraud-Fälle als Hypothesen in Form einer Fall-Datenbank erfasst und in Datensätzen wiedergefunden werden (siehe Arbeitspaket G3), was jedoch aufgrund der eingeschränkten Performanz und der besprochenen Herausforderungen bei der Modellierung nicht in der Praxis umgesetzt werden konnte. Eine entsprechend detaillierte Hypothesenmodellierung ermöglicht gleichzeitig die Anwendung von Regel- und Baumbasierten Algorithmen zur Manipulationserkennung mit dem Vorteil, direkt Aussagen und Vorhersagen über konkrete Fälle treffen zu können, sodass die Untersuchung mit

hypothesenbasierten Ansätzen keinen methodischen Mehrwert für die Manipulationserkennung bieten.

5 Teilarbeitspaket E4: Generalisierbarkeit von Hypothesenbasierten Ansätzen

Wie in D4 untersucht, ergibt sich eine Generalisierbarkeit auch für hypothesenbasierte Ansätze zur Manipulationserkennung durch die abstrakte Modellierung der entsprechenden Datenbasis und die Kompatibilität der Hypothesen.

Im Rahmen dieses APs wurden durch Agentenbasierte Simulation verschiedene Fraud Szenarien in unterschiedlichen Business Settings mit entsprechenden Normal-, Betrugs- und Datenhypothesen simuliert. Die aus diesen Simulationen erzeugten Daten wurden dann in unterschiedlichen Detailgraden als Prozessgraph modelliert, wobei die gemeinsame Fraud Hypothese über die Datensätze hinweg durch Austausch der jeweiligen Modellierungsschichten generalisiert wurde. Hierdurch konnte gezeigt werden, dass unter der Annahme, dass ein Fraud Fall geeignet über Datensatz und Detailierungsgrade hinweg als Hypothese formuliert werden kann, Generalisierbarkeit grundsätzlich gegeben ist, d.h. Evidenz für das Vorkommen entsprechender Fälle durch Hypothesenbasierte Methoden in anderen Datensätzen erfolgreich untersucht werden kann.

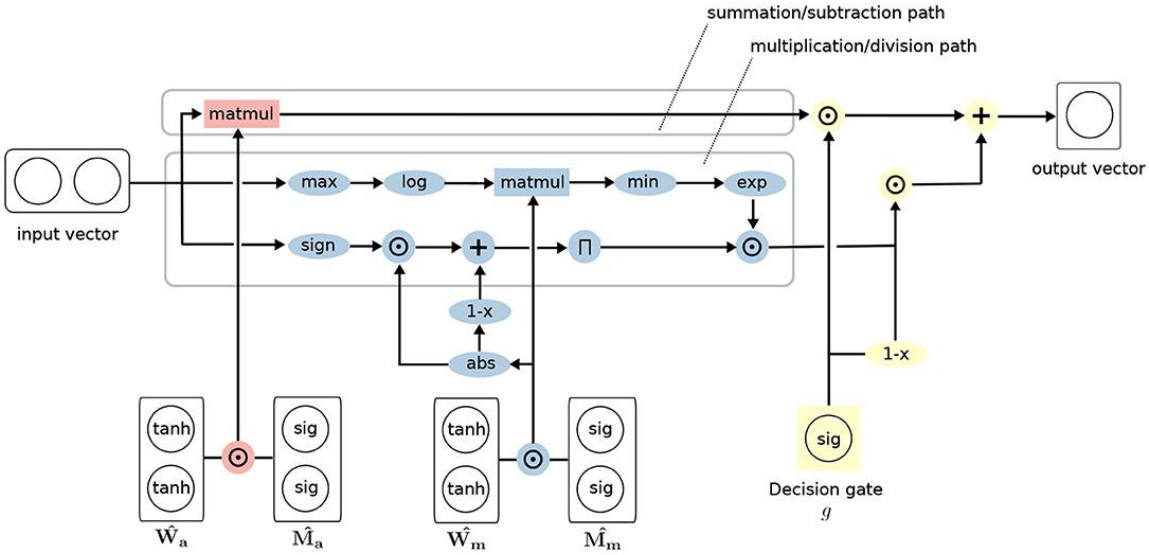

Abbildung 7: iNALU Architecture (Schlör, 2020a)

Die Annahme, dass eine entsprechende Formulierung entsprechend der hier untersuchten synthetischen Simulation auch in der Praxis möglich und sinnvoll ist, ist jedoch wie in E3 diskutiert, nicht gegeben. Dadurch ergibt sich für die folgenden APs ein Fokus auf Anomaly Detection basierte (End-to-end) Ansätze, die sich als für die Praxis geeignetere Wahl und damit vielversprechendere Methodik zur Entwicklung der ML-Toolbox darstellt.

6 Teilarbeitspaket E5: End-to-end Ansätze und Erklärbarkeit

Neben den in Arbeitspaketen E3 entwickelten Verfahren wurden zusätzlich Verfahren evaluiert, welche direkt auf den Rohdaten Manipulationen erkennen können (end-to-end).

Zunächst wurde die Modellierung von kategorialen und arithmetischen Beziehungen untersucht, welche als Kernanforderung identifiziert wurden. Die Umsetzung dieser Beziehungen wurde im diskriminativen und generativen Setting mit Hilfe eines verbesserten Wasserstein GAN untersucht.

Für eine bessere Modellierung von arithmetischen Beziehungen wurde dann die Nutzung von Neural Arithmetric Logic Units (NALU) untersucht, welche allerdings keine ausreichende Genauigkeit auf Daten der finanziellen Betrugs- und Manipulationserkennung lieferte.

Um eine verbesserte Modellierung von arithmetischen Beziehungen zu ermöglichen, wurde deshalb mit der *improved Neural Arithmetic Logic Unit* (iNALU) eine Weiterentwicklung der NALU zur Verbesserung der arithmetischen Genauigkeit und Stabilität entwickelt **(vgl. Abb. 2; Schlör, 2020a)**.

Anschließend wurde eine Testumgebung für die Evaluierung von Methoden der Betrugserkennung auf verschiedenen verfügbaren Finanz-Datensätzen entworfen, um End-to-end Ansätze auf ihre generelle Eignung zur Manipulationserkennung zu untersuchen.

Hierbei wurden insbesondere die im Arbeitspaket D2 zur semantischen Einbettung erstellten Autoencoder-Netzwerke untersucht, da diese auch für eine direkte Erkennung von Manipulationen genutzt werden können.

Des Weiteren wurde auch die entwickelte iNALU Verbesserung für die Anwendung im Rahmen der Manipulationserkennung angepasst, indem eine spezielle Autoencoder Architektur mit einer Mischung aus herkömmlich verwendeten Aktivierungsfunktionen und den hierfür entwickelten iNALU Aktivierungen (Mixed-Layer Architektur vgl. Abb. 3) erstellt wurde **(Schlör, 2020b)**. Die Eignung der resultierenden Architektur wurde auf verschiedenen Finanzdatensätzen gezeigt.

In einem weiteren Schritt wurde eine Testumgebung für End-to-end Ansätze des maschinellen Lernens auf Manipulationsversuchen in ERP-Daten erstellt, welche ausführliche Evaluationen unterschiedlicher Algorithmen und Datenvorverarbeitungsschritte auf ERP-Daten ermöglicht.

Die Autoencoder-basierten Ansätze aus Arbeitspaket D2, sowie die im Projekt entwickelte Mixed-Layer Architektur wurden hiermit auf ihre Performanz bei der Erkennung von Manipulationen in ERP-Daten evaluiert.

Neben Autoencoder-basierten Ansätzen wurden mit Hilfe der erstellten Testumgebung zusätzlich mehrere etablierte Methoden der Anomalieerkennung auf ihre Eignung zur Erkennung von Manipulationen in ERP Daten überprüft.

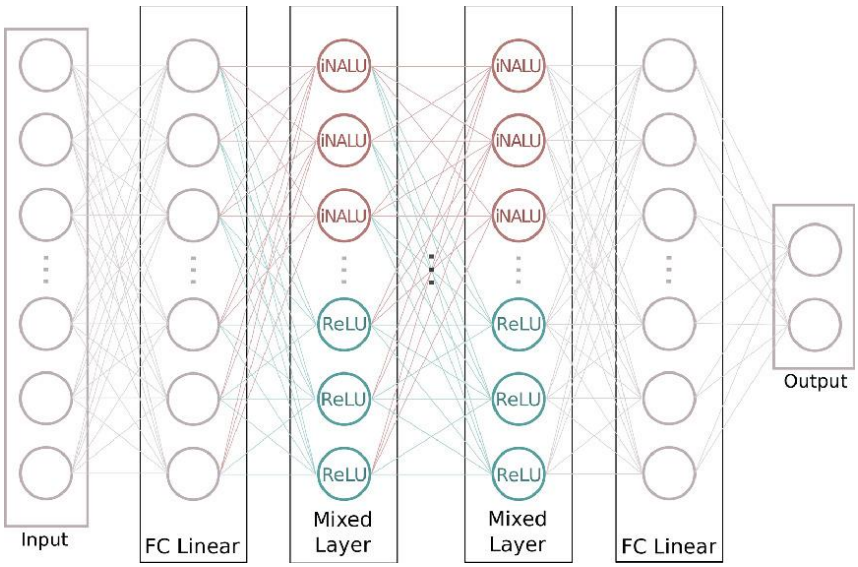

Abbildung 8Mixed-Layer Netzwerk (Schlör, 2020b)

Insgesamt konnten mit End-to-end Ansätzen im Vergleich zu hypothesenbasierten Ansätzen und distanzbasierten Erkennungsmethoden deutlich höhere Detektionsraten erzielt werden, wobei insbesondere Autoencoder Architekturen aus Arbeitspaket D2, die Mixed-Layer Architektur mit iNALU Aktivierungen, sowie ein etablierter Ansatz der Anomalieerkennung (One-Class Support Vektor Maschine) gute Ergebnisse erzielten.

Während die evaluierten End-to-end Ansätze gute Performanz bei der Erkennung von Manipulationen zeigten, besitzen diese jedoch einen undurchsichtigen und nicht nachvollziehbaren Entscheidungsprozess, welcher sie als undurchsichtige Black-Boxen erscheinen lässt.

Da diese Undurchsichtigkeit die Anwendung in der Praxis einschränken kann, wurde auf aktuelle Forschung zur erklärbaren künstlichen Intelligenz

zurückgegriffen und untersucht, ob mit Hilfe dieser Verfahren eine erklärbare Entscheidungsfindung der performanten End-to-end Modelle gewährleistet werden kann.

In einer ersten Arbeit wurde die Genauigkeit verschiedener Erklärungsverfahren aus dem Bereich der post-hoc Feature Relevanz auf ihre Eignung

bei der Anwendung auf kategorische Daten untersucht. Hierfür wurde ein Verfahren entwickelt, um synthetische Benchmarkdaten unterschiedlicher Komplexität zu generieren. Verschiedene State-of-the-art XAI Methoden wurden schließlich in diesem Framework systematisch evaluiert, um die Eignung im Kontext von ERP-Systemen festzustellen **(Tritscher, 2020)**.

In einem weiteren Schritt wurde das hieraus identifizierte Verfahren mit bester Eignung auf kategorische Daten zur Erklärung der End-to-end Ansätze bei der Manipulationserkennung in ERP-Systemen angewendet. Für eine umfangreiche Evaluierung der Erklärbarkeit der unterschiedlichen End-to-end Ansätze, wurde ein Evaluationsschema entworfen und die verschiedenen Ansätze auf qualitative, quantitative und Konsistenzkriterien untersucht **(Tritscher,**

2022b). Hier zeigte sich insbesondere eine sehr gute Interpretierbarkeit der Mixed-Layer Architektur mit iNALU Aktivierungen.

7 Zusammenfassung der Ergebnisse

Zur Übersicht werden im Folgenden die Ergebnisse des Arbeitspakets E bzw. dessen Teilarbeitspakete AP E1 bis E5 für das Forschungsprojekt zusammengefasst.

Basierend auf Experteninterviews und Literaturrecherche konnten im ersten Schritt Anforderungen für die in diesem Arbeitspaket entwickelten Ansätze strukturiert erfasst werden.

Die Evaluation verschiedener Distanzmaße und Metric Learning Ansätze für die Manipulationserkennung in ERP-System Daten ergab teilweise eine akzeptable Leistung bei der Erkennung einzelner Manipulationsfälle. Eine zuverlässige Erkennung *aller* untersuchter Manipulationsfälle konnte durch diese Ansätze allerdings nicht erreicht werden.

Bei der Untersuchung hypothesenbasierter Ansätze für die Betrugserkennung stellte sich die Modellierung der Anomalien als besondere Herausforderung heraus, was sowohl empirisch als auch theoretisch problematisiert wurde. Insgesamt zeigte sich, dass in den durchgeführten Experimenten die hypothesenbasierten Ansätze keinen methodischen Mehrwert für die Manipulationserkennung bieten.

Bei Untersuchung der Generalisierbarkeit hypothesenbasierter Ansätze konnte für synthetischen Datensätze über Datensatz und Detaillierungsgrade hinweg festgestellt werden, dass Generalisierbarkeit gegeben ist, wenn Hypothesen kompatibel formuliert werden können.

Für die untersuchten End-to-end Verfahren zur Manipulationserkennung wurde zunächst eine neuronale Netz-Architektur zur Modellierung von numerischen Zusammenhängen entwickelt, welche anschließend zusammen mit mehreren etablierten Verfahren auf ihre Eignung zur Erkennung von Manipulationen evaluiert wurde. Dabei wurden mit dieser Architektur gute Ergebnisse zur Erkennung von Manipulationen erzielt. Des

Weiteren wurden umfangreiche Evaluationen zur post-hoc Erklärung der performanten Modelle durchgeführt. Mittels einer eigens entwickelten Evaluationsmethodik wurden hierbei Modelle identifiziert, die sowohl gute Erkennungsleistung als auch gute Nachvollziehbarkeit für den Einsatz in der Praxis bieten.

8 Literaturverzeichnis[5]

- *Amarbayasgalan, Tsatsral, Bilguun Jargalsaikhan, and Keun Ho Ryu. "Unsupervised novelty detection using deep autoencoders with density based clustering." Applied Sciences 8.9 (2018): 1468.*
- *Baader, G., & Krcmar, H. (2018). Reducing false positives in fraud detection: Combining the red flag approach with process mining. International Journal of Accounting Information Systems, 31, 1-16*
- *Bartkowiak, Anna M. "Anomaly, novelty, one-class classification: a comprehensive introduction." International Journal of Computer Information Systems and Industrial Management Applications 3.1 (2011): 61-71.*
- *Cirqueira, D., Helfert, M., & Bezbradica, M. (2021, July). Towards design principles for user-centric explainable AI in fraud detection. In International Conference on Human-Computer Interaction (pp. 21-40). Springer, Cham.*
- *De Vazelhes, William, et al. "metric-learn: Metric Learning Algorithms in Python." J. Mach. Learn. Res. 21 (2020): 138-1.*
- *Domingues, Rémi. "Probabilistic Modeling for Novelty Detection with Applications to Fraud Identification." arXiv preprint arXiv:1903.01730 (2019).*
- ***Fuchs, Anna, et al.** "A Meta-Model for Real-Time Fraud Detection in ERP Systems." 5 Jan. 2021, scholarspace.manoa.hawaii.edu/items/2b465ec7-18b1-4fc9-a167-e676c7ed9314/full.*
- *Oosterlinck, Dieter, Dries F. Benoit, and Philippe Baecke. "From*

[5] Eigene Publikationen wurden fettgedruckt gekennzeichnet

one-class to two-class classification by incorporating expert knowledge: Novelty detection in human behaviour." European Journal of Operational Research 282.3 (2020): 1011-1024.

- ***Schlör, Daniel, et al.*** *"Financial Fraud Detection with Improved Neural Arithmetic Logic Units." Workshop on Mining Data for Financial Applications. Springer, Cham, 2020b.*
- ***Schlör, Daniel, Markus Ring, and Andreas Hotho.*** *"inalu: Improved neural arithmetic logic unit." Frontiers in Artificial Intelligence 3 (2020a): 71.*
- ***Tritscher, Julian, et al.*** *"Open ERP System Data For Occupational Fraud Detection." arXiv preprint arXiv:2206.04460, 2022a.*
- ***Tritscher, Julian, et al.*** *"Evaluation of post-hoc XAI approaches through synthetic tabular data." International symposium on methodologies for intelligent systems. Springer, Cham, 2020.*
- ***Tritscher, Julian, et al.*** *"Towards Explainable Occupational Fraud Detection." Workshop on MIning DAta for financial applicationS (MIDAS) . Springer, 2022b.*

Arbeitspaket F: Dynamik der Systeme und Effizienz der Ansätze

Überblick:

In Arbeitspaket F wurden, die in vorherigen Arbeitspaketen entwickelten, Algorithmen zur Manipulationserkennung in ERP-Systemen auf für den Praxiseinsatz relevante Kriterien untersucht und dahingehend verbessert. Ansätze wurden auf ihre Effizienz und ihre Eignung bei der dynamischen Änderung der Datenlage untersucht. Zusätzlich wurde eine Erweiterung der entwickelten Ansätze auf ein Active-Learning-Szenario evaluiert.

Inhaltsverzeichnis Arbeitspaket F

1 Methodik, Vorgehensweise und Unterteilung in Teilarbeitspakete

Arbeitspaket F widmete sich der Optimierung der in Arbeitspaketen D und E entwickelten Verfahren zur Anomalieerkennung. Dabei standen im Wesentlichen drei Teilgebiete im Fokus: Die Anwendung in Echtzeitszenarien, die Robustheit der Verfahren bei Datenänderungen und der Einsatz in einem sogenannten Active Learning Scenario.
In Teilarbeitspaket F1 wurden für alle drei im Arbeitspaket F fokussierte Teilgebiete nötige Anforderungen und zur Bewertung benötigte Kriterien mittels quantitativer Analysen und qualitativer Interviews erarbeitet.

Anschließend wurden in Teilarbeitspaket F2 die in AP D und E entwickelten Ansätze, sowie die dafür in der Praxis benötigten Datenvorverarbeitungs- und Selektionsschritte mittels eines großen Echtdatensatzes auf ihre Effizienz geprüft und notwendige Anpassungen zur Optimierung der Laufzeit durchgeführt.

Um die Robustheit der Algorithmen bei Änderungen der Datengrundlage zu untersuchen, wurden die entwickelten Algorithmen in Teilarbeitspaket F3 mit Hilfe mehrerer Datensätze mit unterschiedlichen Firmenstrategien und wirtschaftlichen Szenarien auf das Beibehalten ihrer Performanz getestet.

Abschließend wurde in Teilarbeitspaket F4 eine Erweiterung der entwickelten Ansätze auf ein Active-Learning-Szenario evaluiert. Hierfür wurde in einem ersten Schritt eine für Betrugserkennung in großen Datenmengen geeignete Annotationspipeline entwickelt, um schnelle Expertenannotationen von Datenpunkten im Active-Learning Kontext zu ermöglichen. Anschließend wurde in einem zweiten Schritt ein modulares Active-Learning Framework für die Anwendung in der Betrugserkennung konzipiert und umgesetzt.

2 Teilarbeitspaket F1: Anforderungsanalyse

Zur Anforderungsanalyse in AP F1 wurden einerseits quantitative

Analysen von Eventlogs und Transaktionsdaten und andererseits qualitative Interviews mit möglichen Anwendern aus der Praxis (Wirtschaftsprüfern, Finanzanalysten, IT-Security) durchgeführt.

Für die Echtzeiteffizienz und Robustheit wurden insgesamt sechs Datensätze aus unterschiedlichen Systemen auf ihre Transaktionsmengen und deren Varianz in enthaltenen Features hin untersucht. Daraus resultierend wurde eine Einschätzung der im Schnitt pro Jahr bearbeiteten Datenmenge erstellt und die Anforderungen zur Effizienz der Systeme hiermit formalisiert. Resultate zeigten einen Belegumfang von ca. 650 Belegen pro Tag, wobei eine starke Varianz der Belegmenge abhängig von Branche, Geschäftsstrategie und Zeit beobachtet werden konnte.

Für das Active Learning Szenario wurden basierend auf dem Design Science Research Ansatz zuerst Design Requirements für ein Anomalieerkennungssystem gesammelt (vgl. Fuchs, 2021) und dann mittels einer Conjoint-Analyse mit insgesamt 18 Probanden basierend auf Eigenschaften aus Wanner et al. (2022) die Erklärungsverfahren zur nachvollziehbaren Visualisierung von Betrugsfällen verglichen.

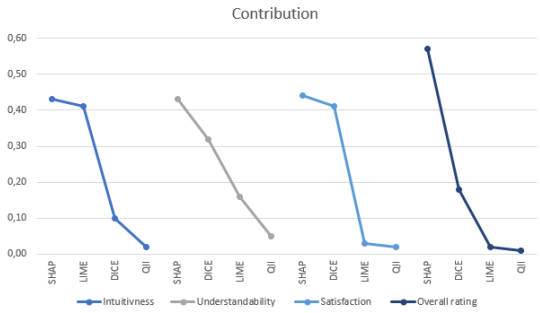

Abbildung.1 Ergebnisse der Conjoint-Analyse

Zwischen den Post-hoc Explainer Verfahren SHAP, LIME, DICE und QII stellte sich bei den Anwendern SHAP als das präferierte Verfahren zur Erklärung der Daten heraus (Ribeiro et al., 2016; Lundberg et al., 2017; Mothilal et al., 2020; Datta et al., 2016).

3 Teilarbeitspaket F2: Steigerung der Effizienz der Algorithmen

Um die entwickelten Ansätze so effizient zu gestalten, dass sie in einem Produktivsystem in Echtzeit arbeiten können, wurde sich im Folgenden auf

die in Arbeitspaket E als geeignet herausgearbeiteten Ansätze fokussiert.
Für die Bewertung und schließlich Steigerung der Effizienz konnte im Rahmen des Projekts ein großer anonymisierter SAP Echtdatensatz gewonnen werden, der zwar aufgrund fehlender bekannter Betrugsfälle nicht für eine Evaluation der Anomalieentdeckung, wohl aber zur Bewertung der Performanz der Ansätze und Algorithmen geeignet war. Zur Anwendung in diesem Big Data Kontext, wurden die zuvor auf kleineren Datensätzen konzipierten Selektions- und Vorverarbeitungsschritte evaluiert und unter Zuhilfenahme von Big-Data Frameworks für die großen Datenmengen angepasst.

Als hoch performante Ansätze, die aus AP E hervorgehen, haben sich speziell die Neuronalen Netze herausgestellt, die nach erfolgtem Training gut parallelisierbar mit sehr geringer Laufzeit auf Transaktionsebene angewendet werden können. Zusammen mit den in diesem AP erfolgten Anpassungen für große Datenmengen zeigen diese eine ausreichende Effizienz im Sinne der Kriterien aus AP F1.

4 Teilarbeitspaket F3: Sicherstellung der Robustheit bei Änderungen

Neben einer ausreichenden Effizienz, stellt auch die Robustheit der Algorithmen gegen Änderungen der Datenbasis einen wichtigen Aspekt für den Einsatz der Algorithmen in der Praxis dar. Änderungen können durch neues Trainieren der Algorithmen auf der Datenbasis realisiert werden. Die vollständige Wiederholung des Trainingsprozesses zieht allerdings teure Evaluationsschritte zum Finden der besten Architekturparameter mit sich, für die manuell von Experten gelabelte Datensätze erstellt werden müssen.
Um wiederholt anfallende teure Evaluationsschritte in der Praxis zu vermeiden, wurden in diesem Arbeitspaket die entwickelten Algorithmen auf ihre Robustheit bei neuen Produkten sowie Änderung von Produktion und genereller Firmenstrategie untersucht.
Für Evaluierungen der Robustheit der Ansätze aus AP E wurden die in AP C5 erstellten Datensätze genutzt. Hierbei wurden Experimente

durchgeführt, um zu evaluieren, ob auf einem Datensatz eines simulierten Unternehmens gewonnene Erkenntnisse über die Leistungsfähigkeit der Ansätze sich auf weitere Datensätze mit unterschiedlichen wirtschaftlichen Szenarien und Handlungsstrategien übertragen lassen. Die Algorithmen wurden auf diesen Datensätzen hinsichtlich ihrer Erkennungsgenauigkeit und der Erklärbarkeit ihrer Entscheidungen evaluiert.
Auswertungen zeigten hier insbesondere sehr robustes Verhalten der in AP E5 erstellten Mixed-Layer Architektur. Resultate der Evaluierung wurden in einem wissenschaftlichen Konferenzbeitrag publiziert (Tritscher et al., 2022).

5 Teilarbeitspaket F4: Evaluation eines Active Learning Szenarios

Um die entwickelten Systeme kontinuierlich zu verbessern, wurde in diesem Arbeitspaket eine Annotationspipeline für ein Active Learning Szenario entwickelt. Für die Umsetzung von Active Learning in diesem Setting stehen zwei Herausforderungen im Mittelpunkt: Zum einen müssen dem Annotator die richtigen Daten angezeigt werden, die geeignet sind, um das Fraud Detection Modell zu verbessern, zum anderen müssen die zu bewertende Datenpunkte adäquat präsentiert werden, sodass der Annotator seine Bewertung schnell und präzise vornehmen kann.
Bei der Präsentation der Datenpunkte ergibt sich insbesondere durch das im Projekt bearbeitete Szenario der Anomalie- und Betrugserkennung die Notwendigkeit, neben einzelnen Datenpunkten zusätzlich das normale Regelverhalten der Daten zu repräsentieren und so dem Annotator eine Abgrenzung des zu untersuchenden Datenpunktes zum Regelverhalten zu ermöglichen.

Um dem Annotator sowohl den zu untersuchenden Datenpunkt als auch das Regelverhalten angemessen zu präsentieren, wurden basierend auf aktuellen Erkenntnissen der visuellen Anomalieerkennung sowie der interaktiven Visualisierung relevante Visualisierungstechniken identifiziert und eine passende Annotationsoberfläche entwickelt.

Zusätzlich wurden bei der Darstellung des Regelverhaltens für die Betrugserkennung in ERP-Systemen große Datenmengen als Herausforderung identifiziert, da anfallende Datenmengen sowohl eine direkte Visualisierung der Gesamtdatenmenge als auch die exakte Berechnung von statistischen Kenngrößen behindern.

Weiterhin wurde die Eignung von Samplingverfahren untersucht, um das Regelverhalten innerhalb angemessener Zeit visuell zu repräsentieren. Dazu wurden mehrere Samplingverfahren dahingehend untersucht, ob gezogene Stichproben der Daten adäquat das Regelverhalten des Gesamtdatensatzes widerspiegeln. Insbesondere zeitbasierte Samplingverfahren, welche ganze Tage und Wochen als Stichproben extrahieren, wurden für diese Anwendung als geeignet befunden.

Basierend auf den entwickelten Teilen zur Auswahl der Datenpunkte und zur Visualisierung wurde dann ein Active Learning Fraud Detection Framework konzipiert. Zentral für die modulare Architektur ist die Trennung von Datenhaltung, Fraud Detection Modellinstanzen, Active Learning Strategie und Visualisieren, die für jedes Projekt einzeln ausgewählt und konfiguriert werden können. Hierdurch konnte den unterschiedlichen Anforderungen von überwachten und unüberwachten Verfahren begegnet werden, sowie ein für jeden Auditor und Datensatz spezifisches Dashboard zur effizienten Kontrastierung von potenziellen Fraud- und Normaldaten realisiert werden.
Die gemeinsame Evaluation mit Anwendern aus Unternehmen vor Ort konnte bis zum Projektabschluss COVID-19 bedingt nicht durchgeführt werden. Die hierfür notwendigen Komponenten sind jedoch vorbereitet und sollen, soweit möglich, im Nachgang des Projekts evaluiert werden.

6 Zusammenfassung der Ergebnisse

Im Rahmen des Arbeitspakets F wurden die entwickelten Algorithmen auf die Problemstellungen der Effizienz und Dynamik untersucht.
Hierbei wurde in Teilarbeitspaket F2 die Effizienz der Algorithmen und

benötigten Datenvorverarbeitungs- und Selektionsschritte durch Experimente und Anpassungen für einen Big Data Echtdatensatz sichergestellt.

Die Robustheit der Algorithmen gegenüber variierender Datencharakteristiken und wirtschaftlicher Szenarien wurde in Teilarbeitspaket F3 mit Experimenten auf Datensätzen aus C5 evaluiert und erfolgreich robuste Ansätze identifiziert.

Für die Erweiterung der Ansätze mittels Active-Learning wurde in Teilarbeitspaket ein Annotationstool entwickelt, Experimente zur Nutzbarkeit der Annotationsumgebung im Big Data Kontext durchgeführt, sowie ein modulares Active-Learning Framework für die Erkennung von Manipulationen in ERP-Systemen konzipiert.

7 Literaturverzeichnis[6]

- *Datta, Anupam, Shayak Sen, and Yair Zick. "Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems." 2016 IEEE symposium on security and privacy (SP). IEEE, 2016.*
- *Lundberg, Scott M., and Su-In Lee. "A unified approach to interpreting model predictions." Advances in neural information processing systems 30 (2017).*
- *Mothilal, Ramaravind K., Amit Sharma, and Chenhao Tan. "Explaining machine learning classifiers through diverse counterfactual explanations." Proceedings of the 2020 conference on fairness, accountability, and transparency. 2020.*
- *Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. "Why should i trust you? Explaining the predictions of any classifier." Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016.*
- ***Tritscher, Julian, et al.** "Towards Explainable Occupational Fraud Detection." Workshop on MIning DAta for financial applicationS (MIDAS) - forthcoming. Springer, 2022.*
- ***Wanner, Jonas, et al.** "A social evaluation of the perceived goodness of explainability in machine learning." Journal of Business Analytics, vol. 5, no. 1, 2 Jan. 2022, pp. 29-50, doi:10.1080/2573234X.2021.1952913.*

[6] Eigene Publikationen wurden fettgedruckt gekennzeichnet

Arbeitspaket G: Implementierung und Evaluation in der Praxis

Überblick:

Arbeitspaket G beinhaltet die direkte Integration der in vorherigen Arbeitspaketen entwickelten Algorithmen in das ERP-System.
Hierbei wurde eine prototypische Umsetzung einer Machine Learning Toolbox als ERP-Erweiterung mit generalisierbaren Schnittstellen entwickelt und diese mit Hilfe der ERP-Systeme der Konsortialpartner auf Einsatzfähigkeit in Theorie und Praxis evaluiert.

Inhaltsverzeichnis Arbeitspaket G

1 Vorgehensweise, Unterteilung in Teilarbeitspakete und Methodik

Nach der Entwicklung von Methoden zur Entdeckung von Manipulationen in Arbeitspaketen D und E, sowie der Optimierung der Verfahren in Arbeitspaket F, steht in Arbeitspaket G die Adaption der Verfahren im Fokus.

Dazu wurden im ersten Schritt (G1) die sozioökonomischen, technischen und wirtschaftlichen Rahmenbedingungen analysiert und relevante Anspruchsgruppen identifiziert. Basierend auf den abgeleiteten Rahmenbedingungen und technischen Anforderungen aus Arbeitspaket C wurden eine prototypische Umsetzung des ERP-Add-Ins (G2) sowie eine Machine Learning Toolbox realisiert (G3). Die Anbindung der Komponenten erfolgte in Teilarbeitspaket G4, wobei die Einhaltung der datenschutzrechtlichen Anforderungen (G5) nochmals explizit Beachtung fand.

Methodisch wurde für G1 eine weitere Anforderungsanalyse (Requirements Engineering) mit relevanten Interessengruppen durchgeführt. In den folgenden Arbeitspaketen wurden basierend auf den gesamten Anforderungen eine entsprechende prototypische Umsetzung entwickelt.

2 Teilarbeitspaket G1: Anforderungsanalyse

Der Fokus der Anforderungsanalyse zu Arbeitspaket G lag einerseits auf dem Gesamtkonstrukt einer Anwendung, der technischen Anbindung und den nötigen Anpassungsmöglichkeiten, sowie andererseits auf der Generalisierung der Verfahren im Sinne einer Machine Learning Toolbox.
Die Anforderungsanalyse bestand aus einer Design Science Studie (vgl. Fuchs, 2021) und der Erweiterung der Anforderungen bzgl. der Erklärbarkeit nach Cirqueira et al. (2021).
Außerdem wurden in einer ersten sozioökonomischen Überblicksanalyse, relevante Anspruchsgruppen herausgearbeitet.

Tabelle 5: Ermittelte Anforderungen

ID	Anforderungen
R1	Unterstützung komplexer Fraudszenarien
R2	Analyse großer Datenmengen
R3	Erkennen von Verfahrensänderungen & Prozessabweichung
R4	Unterstützung von Dokumentenfälschung
R5	Erkennung von Ausreißern in den Werten
R6	Integration in alle ERP-Funktionen / Module
R7	Erkennung von Datendiebstahl
R8	Unmittelbare Betrugserkennung
R9	Unterstützung von unbekannten Szenarien zusätzlich zur Mustererkennung oder Prüfsummen
R10	Lernfähig, anpassungsfähig & intelligente Logik
R11	Integration der menschlichen Erfahrung für ein besseres und verständlicheres System
R12	Bereitstellung vorhergesagter Fraudfälle und der Einschränkungen für die Klassifizierung von Fällen
R13	Bereitstellung ähnlicher und unähnlicher klassifizierter Fraudfälle
R14	Bereitstellung einer dynamische Sicht auf Daten und Details
R15	Bereitstellung von Beziehungen zwischen Attributen in einzelnen und mehreren klassifizierten Fraudfällen
R16	Bereitstellung der Bedeutung der vom KI-Modell verwendeten Attribute für die Klassifizierung von Fraudfällen
R17	Bereitstellung des Argumentationsprozess des KI-Modells zur Klassifizierung von Fraudfällen
R18	Bereitstellung der Auswirkungen der vom KI-Modell verwendeten Attribute auf bestimmte Klassifizierungen von Fraudverdachtsfällen

Dazu wurde Kontakt zu Unternehmen unterschiedlichster Branchen (u.A. Schwarz AG, Audi, Deloitte, SAP) hergestellt und der Einsatz der konzipierten Lösung bzw. relevante Interessensgruppen diskutiert.

Die herausgearbeiteten Anforderungen (vgl. Tabelle 1) wurden außerdem in Workshops mit Anwendern der drei wesentlichen relevanten Anspruchsgruppen, (ERP-) Softwareanbieter, unternehmensexterne Anwender (Wirtschaftsprüfung) und interne Anwender (IT-Controlling / IT-Compliance / Finance), evaluiert und weiter detailliert.

3 Teilarbeitspaket G2: Entwicklung einer ERP-Erweiterung

Für die Umsetzung als ERP-Add-In wurde eine Microservice-basierte Architektur in Form eines Software-Containers gewählt. Hierdurch lässt sich die Anwendung sowohl auf Cloud-Anbietern wie Amazon AWS, Google GCP und Microsoft Azure, aber auch auf eigenen Servern möglichst einfach bereitstellen. Die Entwicklung als Container ist dabei, insbesondere hinsichtlich des Datenschutzes, aber auch der verfügbaren Hardware in Unternehmen, der einfachen Adaptionsfähigkeit zutragend. Je nach Datenmenge und Bedarf können für das Training des neuronalen Netzes speziell dafür entwickelte

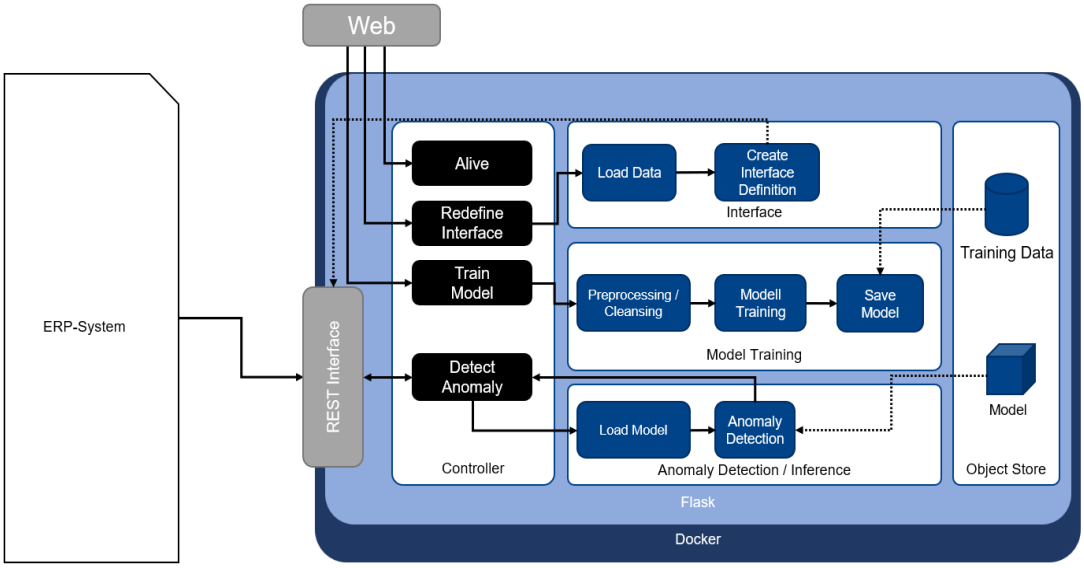

Abbildung 9: Architektur des ERP-Add-Ins

Compute- / GPU-Ressourcen verwendet werden und für den Produktiveinsatz der Anwendung (Anomalieerkennung) sowohl eigene als auch ressourcenschonend externe Infrastruktur genutzt werden.

Das ERP-Add-In besteht im Wesentlichen aus einem Framework Controller, der, basierend auf
Flask, die Zugriffe auf die Applikation mit Hilfe eines Trainings- und Anomalieerkennungs- / Inferencing Moduls steuert. Der Datensatz und das Modell werden dabei nicht versioniert im Object Store innerhalb des Containers gesichert.

Insgesamt wurden dazu vier mögliche Funktionsaufrufe zur simplifizierten Interaktion und Steuerung der Applikation implementiert. Die *alive* Funktion ermöglicht es Service-Monitoring Systemen und Orchestrierungsframeworks wie bspw. Docker-Compose den Status der Applikation abzufragen. Um eine dynamische Anpassung der REST-API auf die von Anwendern zum Training bereitgestellten Datensätze zu ermöglichen, wurde die Funktion *Redefine Interface* angelegt. Die Anwendung bietet außerdem eine Funktion zum (neu-)Trainieren des Modells zur Anomalieerkennung, basierend auf vorab in XML bereitgestellten Trainingsdaten. Das Ergebnis des Trainings für den Trainingsdatensatz lässt sich dabei über eine externe Webapplikation analysieren.
Die Funktion zur Anomalieerkennung auf neuen Datensätzen, die direkt von der Schnittstelle zum ERP-System aufgerufen wird, lädt das vorab trainierte Modell und gibt das Ergebnis direkt über die REST-Schnittstelle als Antwort an das ERP-System zurück.

4 Teilarbeitspaket G3: Realisierung einer Machine Learning Toolbox

Die gesamte Machine Learning Toolbox wurde flexibel entwickelt und besteht aus mehreren Teilen. Als Grundlage für eigene Data-Pipelines oder Softwareprodukte sind die im Projekt entwickelten Algorithmen öffentlich zur Adaption verfügbar. Zusätzlich bietet die Adaptierbarkeit des Add-Ins die Möglichkeit, die entwickelten Algorithmen für jegliche Anwendungen mit entsprechenden Schnittstellen zu übertragen.

Da End-to-end Ansätze auf unterschiedlichen Datenbasen (vgl. D5) im Gegensatz zu hypothesenbasierten Verfahren keine datenschutzkonforme Bereitstellung von abstrakten Fallinformationen ermöglichen und hypothesenbasierte Ansätze aufgrund der vergleichsweise schwächeren Performanz (vgl. E4) in der Toolbox schlussendlich nicht zum Einsatz kamen, hat sich die zentrale Bereitstellung einer Falldatenbank demnach als nicht umsetzbar erwiesen.

Zusätzlich zum ERP-Add-In wurde aber eine Stand-Alone Anwendung als Komponente der Machine-Learning Toolbox entwickelt, welche gleichzeitig als Demonstrator dienen soll. Die Stand-Alone Anwendung bietet die Möglichkeit, auch für Anwender ohne Programmierkenntnisse die im Projekt entwickelten Verfahren auf eigenen Datensätzen testweise anzuwenden.

Der Demonstrator wurde dazu möglichst benutzerfreundlich gestaltet und begleitet den Anwender durch den Prozess vom Hochladen der Daten, über die Anonymisierung, bis hin zur Anomalieerkennung und der Ergebnisanalyse.

5 Teilarbeitspaket G4: Entwicklung generalisierbarer Schnittstellen

Die Anbindung des ERP-Add-Ins erfolgte über den REST-Schnittstellen-Standard, da sich REST bereits in AP C als der am weitesten verbreitete Schnittstellen-Typ herausgestellt hat. Obwohl die Festlegung auf REST eine Einschränkung hinsichtlich der Adaptionsfähigkeit darstellt, gehen wir davon aus, dass REST als die wohl zukunftsträchtigste Schnittstellentechnologie durch die fortschreitende Weiterentwicklung im Softwaremarkt in näherer Zukunft von immer mehr Herstellern übernommen werden wird.

Die Schnittstelle vom ERP-System zum Add-In kann dabei dynamisch an das jeweilige ERP-System bzw. dessen Schnittstellendefinition und damit auch an den verwendeten Kontext angepasst werden. Eine entsprechende Funktionalität zur automatischen Generierung der Schnittstellendefinition stellt der Container als Funktion bereit (vgl. Abb. 1).

Die Antwort vom ERP-Add-In zum ERP-System hingegen ist aktuell relativ einfach ausgestaltet. Der REST-Aufruf der Anomaliedetektion mit entsprechenden Daten bietet eine Schnittstelle zur Rückgabe zweier Variablen zurück, die *Prediction* und den *Abnormalitätswert*. Während die *Prediction* als boolsche Variable das Ergebnis der Anomaliedetektion (Fraud / nicht Fraud) darstellt, gibt der Abnormalitätswert den Grad der Abweichung des Datenpunkts zu normalem Verhalten wieder.

Zur vereinfachten Handhabung der REST-Schnittstelle wurde im Container FAST-API genutzt, was die Dokumentation des REST-API in Form eines aufrufbaren, testbaren Webinterfaces ermöglicht.

6 Teilarbeitspaket G5: Implementierung datenschutzrechtlicher Anforderungen

Da in der aktuellen prototypischen Umsetzung des ERP-Add-Ins existierende Hersteller-spezifische REST-Schnittstellen verwendet wurden, ist die Pseudonymisierung und Anonymisierung eine Aufgabe in der jeweiligen Schnittstellendefinition.

In einer ersten Version des Plugins besteht lediglich die Möglichkeit, datenschutzkritische Attribute, welche über die REST-Schnittstelle übertragen werden, zu verwerfen, also Features bewusst aus dem Datensatz und damit aus der Verarbeitung herauszunehmen.

Damit ist es auch möglich einen Connector zu entwickeln, mit dem einzelne Variablen (Features des Datensatzes), mittels geeigneter Verfahren anonymisiert, pseudonymisiert oder entfernt werden können. Hierdurch kann jegliches Datum anwenderdefiniert vor der Übertragung exkludiert oder datenschutzkonform nach den in Arbeitspaket B erarbeiteten Standards behandelt werden.

7 Teilarbeitspaket G6: Evaluation in Theorie und Praxis

Um eine realitätsgetreue Marktsituation zu erzeugen, werden zum Test und der Evaluation des entwickelten ERP-Add-Ins neben realistischen Angriffsszenarien auch realistische Normaldaten benötigt.

Eine Möglichkeit zur Herstellung einer entsprechenden Testumgebung stellt die Erzeugung von Daten innerhalb einer Simulationsumgebung dar. Dazu stand im Projekt lediglich eine geschlossene ERP-Instanz von SAP S/4 HANA mit einer proprietären Simulationsumgebung von ERP-SIM zur Verfügung (vgl. AP C5). In dieser konnten Angriffe zwar simuliert werden, jedoch aufgrund der fehlenden Schnittstellen und Entwicklungsrechte sowie weiterer proprietärer Beschränkungen des Systems keine Integration mit dem ERP-Add-In ermöglicht werden.
Zur weiteren Evaluation wurde das ERP-Add-In außerdem erfolgreich an das produktive ERP-System der Step Ahead angebunden. Hier stand andererseits keine Simulationsumgebung zur Verfügung, um neben manuell modellierten Angriffen realistische Normaldaten in ausreichender Menge zu erzeugen. Aufgrund dieser Einschränkungen wurde die Evaluation wie folgt zweiteilig durchgeführt:

Im **ersten Schritt** wurde eine quantitative Evaluation mit in SAP erzeugten Testdaten durchgeführt, wobei das ERP-Add-In dabei nicht vollintegriert mit dem System zusammenarbeitete. Die Ergebnisse der quantitativen Evaluation basierend auf den in SAP simulierten Angriffen zeigten dabei deutliche Vorteile der in Arbeitspaket E entwickelten Mixed Layer Architektur im Vergleich zu klassischen Verfahren wie der Support Vector Maschine im One-vs-All Ansatz.
In einem **zweiten Schritt** wurde eine qualitative Evaluation der Machine Learning Toolbox im Produktivsystem der Step Ahead durchgeführt. Dabei konnte der praktische Einsatz innerhalb des Step-Ahead-Systems (STEPS ERP) sogar mit realen Daten getestet, jedoch konnten hier keine simulierten Angriffe durchgeführt werden.

Die Ergebnisse der qualitativen Evaluation im Produktivszenario, basierend demnach auf historischen Echtdaten der Step-Ahead und zeigen, dass sich das ERP-Add-In in einem voll integrierten Aufbau nicht nur einfach auf das System einstellen lässt, sondern auch, dass die Anwendung auf Echtdaten wirtschaftlich interessante Ergebnisse erzeugen kann. Unter ungefähr achttausend analysierten

Transaktionen konnten ca. dreißig auffällige Aktivitäten (red-flags) und acht relevante Anomalien (die vorab unbekannt waren) ausgemacht werden. Insgesamt zeigten die qualitative und die quantitative Evaluation damit ein positives Ergebnis hinsichtlich der Einsetzbarkeit in der Praxis.

8 Zusammenfassung der Ergebnisse

Basierend auf der Anforderungsanalyse in G1 und den vorangegangenen Arbeitspaketen erfolgte in den Arbeitspaketen G2-G5 die Entwicklung des ERP-Add-Ins und des Demonstrators sowie die Evaluation der Komponenten.
Hierzu wurde in AP G2 ein auf einer cloud-fähigen Architektur basierender Microservice zur Kapselung der im Projekt fokussierten Deep Learning Verfahren zur Anomaliedetektion entwickelt. Unter der Voraussetzung einer bestehenden REST-API lässt sich dieser, dank variabler Schnittstellendefinition (G4), auf verschiedene Unternehmensprozesse und ERP-Systeme adaptieren.

Zusätzlich wurde eine Stand-Alone-Applikation entwickelt, welche die Anwendung der Anomaliedetektion ohne Schnittstelle und für Anwender ohne Programmierkenntnisse ermöglicht.

Die Evaluation der Komponenten wurde quantitativ und qualitativ angegangen, wodurch nicht nur auf simulierten Angriffen, sondern auch auf Echtdaten erste positive Ergebnisse gezeigt werden konnten.

9 Literaturverzeichnis[7]

- ***Fuchs, Anna, et al.*** *"A Meta-Model for Real-Time Fraud Detection in ERP-Systems." 5 Jan. 2021, scholarspace.manoa.hawaii.edu/items/2b465ec7-18b1-4fc9-a167-e676c7ed9314/full.*
- *Cirqueira, D., Helfert, M., & Bezbradica, M. (2021, July). Towards design principles for user-centric explainable AI in fraud detection. In International Conference on Human-Computer Interaction (pp. 21-40). Springer, Cham.*

[7] Eigene Publikationen wurden fettgedruckt gekennzeichnet

Arbeitspaket H: Dissemination

Überblick:

Um die Ergebnisse an potenzielle Nutzer sowie andere Interessensparteien zu kommunizieren, wurden als Teil des Projekts diverse Kanäle zur Publikation und Verbreitung der Ergebnisse genutzt. Im Folgenden wird ein Überblick über die Veröffentlichungen und Kommunikation von Ergebnissen des Projekts gegeben.

Inhaltsverzeichnis Arbeitspaket H

1 Fokus und Aufteilung der Publikationen

Die Dissemination teilt sich in **Publikationen** unterschiedlichster Art sowie **Wissenschaftskommunikation** über andere Kanäle (wie, Messen und Veranstaltungen). Insgesamt konnten im Rahmen des Projekts rund 10 Publikationen sowie mehrere Veranstaltungen ausgerichtet bzw. im Rahmen von Gastvorträgen Teilaspekte des Projekts vorgestellt werden. Details zu den Beiträgen finden sich auch über die Projekthomepage (https://projekt-deepscan.de/). Aus dem Projekt heraus wurden außerdem diverse studentische Arbeiten vergeben, welche im Folgenden aufgelistet werden.

Des Weiteren werden im Folgenden Arbeiten, die teils über das Projektende hinaus erfolgen, im Sinne der Anschlussfähigkeit aufgeführt.

2 Wissenschaftliche Publikationen

Die folgenden, bereits in den entsprechenden Arbeitspaketen diskutierten, wissenschaftliche Beiträge gingen aus dem Forschungsprojekt DeepScan hervor:

- *Fuchs, A., Fuchs, K., Gwinner, F., & Winkelmann, A. (2021). A Meta-Model for Real-Time Fraud Detection in ERP Systems. Proceedings of the 54th Hawaii International Conference on System Sciences. doi:10.24251/HICSS.2021.856*
- *Hofmann, A., Gwinner, F., Fuchs, K., & Winkelmann, A. (2020). An Industry-Agnostic Approach for the Prediction of Return Shipments. In AMCIS 2020 Proceedings. 32.*
- *Schlör, D. (2023). Detecting Anomalies in Transaction Data. Dissertation - forthcoming.*
- *Schlör, D., Ring, M., & Hotho, A. (2020). inalu: Improved neural arithmetic logic unit. Frontiers in Artificial Intelligence, 3, 71. doi:10.3389/frai.2020.00071.*
- *Schlör, D., Ring, M., Krause, A., & Hotho, A. (2020, September). Financial Fraud Detection with Improved Neural Arithmetic Logic Units. In Workshop on Mining Data for Financial Applications (pp. 40-54). Springer, Cham. doi:10.1007/978-3-030-66981-2_4.*
- *Tritscher, J., Gwinner, F., Schlör, D., Krause, A., & Hotho, A. (2022). Open ERP System Data for Occupational Fraud Detection. arXiv preprint. doi:10.48550/arXiv.2206.04460.*
- *Tritscher, J., Krause, A., Schlör, D., Gwinner, F., Von Mammen, S., & Hotho, A. (2021, August). A financial*

game with opportunities for fraud. In 2021 IEEE Conference on Games (CoG) (pp. 1-5). IEEE. doi:10.1109/CoG52621.2021.9619070.

- *Tritscher, J., Ring, M., Schlr, D., Hettinger, L., & Hotho, A. (2020, September). Evaluation of post-hoc XAI approaches through synthetic tabular data. In International symposium on methodologies for intelligent systems (pp. 422-430). Springer, Cham. doi:10.1007/978-3-030-59491-6_40.*
- *Tritscher, J., Schlör, D., Gwinner, F., Krause, A., & Hotho, A. (2023). Towards Explainable Occupational Fraud Detection. In Workshop on Mining Data for Financial Applications. - forthcoming. Springer, Cham.*

3 Wissenschaftskommunikation

Neben den Publikationen, die sich auf die wissenschaftlichen Ergebnisse und Diskussion fokussieren, wurden im Rahmen des Projekts diverse nichtwissenschaftliche Publikationen erarbeitet, in denen der aktuelle Stand des Projekts an Unternehmen und Industrie kommuniziert wurde. Außerdem gab es verschiedene Beiträge von Medien, bei denen Mitarbeiter zu Gast waren, sowie Veranstaltungen, Messen und Tagungen, an denen teilgenommen wurde.

- *Fuchs, A., Gwinner, F., & Winkelmann, A. Betrug und Manipulationen in Informationssystemen. ERP Management, 16(3),*
- *Projektwebseite: https://projekt-deepscan.de*
- *IHK, WIRTSCHAFT in Mainfranken 08/2019, https://www.wuerzburg.ihk.de/fileadmin/user_upload/WIM/2019/Wirtschaft_in_Mainfranken_August_2019.pdf*
- *Kaminabend der Wirtschaftsinformatik Uni Würzburg, 2018. Vorstellung des Projekts DeepScan.*
- *ESF Veranstaltungsreihe zu betrieblicher Software und Digitalisierung. Projektvorstellung. https://www.uni-wuerzburg.de/sft/esf-promptnet/*
- *IHK, Vortrag beim IT-Sicherheitsforum 2019. https://www.youtube.com/watch?v=Pd84aFq9Lx4*
- *„The Future Code“, Teilnahme und Vorstellung des Projekts mit Messestand. https://www.thefuturecode.de/*
- *FIS Logistics Day 2021: Vortrag zum Thema Anomalieerkennung mittels ML.*
- *ERP des Jahres 2021 / 2022 https://events.gito.de/erp-systeme-des-jahres-2022/*
- *DLF Beitrag, https://www.deutschlandfunk.de/programm?cal:month=1&cal:year=2019&drsearch:date=2019-01-19*

4 Studentische Arbeiten

Im Rahmen der Arbeitsgruppe zum Projekt DeepScan wurden diverse studentische Arbeiten vergeben, um Teilaspekte der Arbeitspakete zu beleuchten, literarische Vorarbeiten zu leisten oder den wissenschaftlichen Stand aufzuarbeiten. Folgende Aufstellung soll einen Überblick über die diversen Thesen bieten, die vom Projekt profitierten.

- *Christian Zeiß (2021), „Effiziente Generierung von Event Logs für Process Mining im Zoll- und Außenhandel -Eine konzeptuell analytische Aufbereitung“, Masterthesis.*
- *Daniel Wilhelm (2022), “Sampling-basiertes Annotationstool für Fraud-Erkennung in ERP Systemen“, Bachelorthesis.*
- *Dominik Gils (2022), “Applying Reinforcement Learning to an ERP-System Simulation”, Masterthesis.*
- *Janina Budnik (2022), „Betriebswirtschaftliche Datenanalyse von Beschaffungsdaten, des öffentlichen Sektors“, Masterthesis.*
- *Katharina Vogel (2022), „Graph Embeddings – Methoden und Anwendung“, Bachelorthesis.*
- *Luca Lauer (2022), “Anomaly Detection and the modelling of normality”, Bachelorthesis.*
- *Lukas Hörner (2022), “Enhancing the Data Value with Encodings”. Masterthesis.*
- *Matthias Habereder (2022), „Multi-Agent Based Simulation for Normal and Fraudulent ERP-Data Generation“, Bachelorthesis.*
- *Moritz Lampert (2022), “Privacy of Neural Network Models”, Bachelorthesis.*
- *Nils Wapki (2022), “Comparison of Big Data Preprocessing and Anomaly Detection approaches for ERP Data”, Bachelorthesis.*
- *Pascal Fries (2022), “Fraud Detection using Graph Neural Networks”, Masterthesis.*
- *Robin Gebhardt (2021), “Konzeption und Umsetzung einer Data Pipepline zur Analyse von ERP-Daten“, Masterthesis.*
- *Thomas Mahlmeister (2022), “Domain-specific Anomaly Detection”, Masterthesis.*

5 Wissenschaftliche und technische Anschlussfähigkeit

Ein Anschluss aus wissenschaftlicher Sicht, um Themen wie Active Learning, Federated Learning oder Graph-ML, wäre aus Sicht der Universität zu begrüßen, da entsprechendes Potenzial und Erfolgsaussichten ersichtlich sind. Das Konsortium ist aktuell schon dabei, weitere thematische

Anschlussmöglichkeiten zu prüfen. Dabei stehen Verfahren wie Graph Machine Learning auf den im Projekt (für HypTrails und SubTrails) erzeugten Prozessgraphen, aber auch der Ausbau zu einer alleinstehenden Active Learning Umgebung, sowie der Einsatz von Federated Learning zur Verbesserung der Algorithmen für Cloudsoftwareanbieter zur Diskussion.

Im Rahmen des Projekts kamen neben den Konsortialteilnehmern diverse Kontakte mit Unternehmen zustande, welche zu einer Zusammenarbeit führten bzw. über das Projekt hinaus Anschlussfähigkeiten, für weitere Zusammenarbeit bieten.

So lieferten, neben der Step Ahead, die Unternehmen JOPP, Eagle Burgmann und Weclapp jeweils real Datensätze, die hinsichtlich ihrer Größe und Varianz im Projekt zur Evaluation der Algorithmen herangezogen werden konnten. Während für Jopp und Eagle Burgmann im Wesentlichen die Ergebnisse der Anomalieerkennung und Datenanalyse interessant waren, wird sowohl mit der Weclapp SE und der Step Ahead eine Weiterentwicklung des ERP-Add-Ins fokussiert. Dabei soll ein Open-Core Modell ermöglicht werden, bei dem lediglich proprietäre Softwareteile lizenziert, der Kern aber frei verfügbar bleibt.

Zusätzlich ergab sich bezüglich des zur Datengenerierung entwickelten Serious-Games (Tritscher et al., 2021) die Möglichkeit einer Zusammenarbeit zur Weiterentwicklung. Das Unternehmen Sopra Banking Software hat Interesse das entwickelte Serious-Game als innovativen Ansatz zur Bekämpfung von Fraud und im weiteren Sinne auch KYC - Know Your Customer bzw. KNE - Know your Employee, zur Schulung aber auch zur Entwicklung Datengetriebener Verfahren, zur Machine Learning basierten Anomalieerkennung, einzusetzen. Aktuell werden hierzu erste Ideen ausgetauscht und Möglichkeiten zur Zusammenarbeit evaluiert.

Insgesamt zeigt das Interesse an der im Projekt entwickelten Anomalieerkennung aber auch an anderen Nebenprodukten wie dem Serious-Game diverse Anschlussfähigkeiten und damit auch den Erfolg des Projekts.

Konsortialpartner

Projektbeirat

Autoren und Ansprechpartner

Prof. Axel Winkelmann
Lehrstuhl für BWL und Wirtschaftsinformatik
axel.winkelmann@uni-wuerzburg.de
+49 931 31-89640

Prof. Andreas Hotho
Lehrstuhl Informatik XI
hotho@informatik.uni-wuerzburg.de
+49 931 31-88453

Fabian Gwinner
Lehrstuhl für BWL und Wirtschaftsinformatik
fabian.gwinner@uni-wuerzburg.de
+49 931 31-81946

Kevin Fuchs
Lehrstuhl für BWL und Wirtschaftsinformatik
kevin.fuchs@uni-wuerzburg.de
+49 931 31-88034

Julian Tritscher
Lehrstuhl Informatik XI
julian.tritscher@uni-wuerzburg.de
+49 931 31-84467

Daniel Schlör
Lehrstuhl Informatik XI
daniel.schloer@informatik.uni-wuerzburg.de
+49 931 31-84564

Die Ergebnisaufbereitung wurde durchgeführt von:

Julius-Maximilians-Universität Würzburg
Sanderring 2
97070 Würzburg

[Stand August 2022, aktuelle Informationen zum Projekt http://projekt-deepscan.de/]

Förderung & Betreuung

Betreut vom

DLR Projektträger

Gefördert vom

Bundesministerium
für Bildung
und Forschung