

Fakultät für
Informatik und
Mathematik



Performance Evaluation Mobiler Anwendungen

Görkem Bas

Bachelorarbeit Informatik

Prüfer:

Prof. Dr. Gudrun Socher, Hochschule München

Betreuer:

Philipp Kapinos, Brunata Metrona GmbH

01.03.2023

Erklärung

Görkem Bas, geb. 24.03.2000 (IF7, WS 22/2023)

Hiermit erkläre ich, dass ich die Bachelorarbeit selbständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

München, 01.03.2023

.....

Unterschrift

Zusammenfassung

Das Ziel der vorliegenden Arbeit ist es zu beantworten, wie man die Performance Mobiler Anwendungen Evaluieren kann. Dazu wird die Folgende Forschungsfrage gestellt: Wie kann man

Danksagungen

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Bachelorarbeit unterstützt und motiviert haben.

Zuerst gebührt mein Dank Frau Prof. Gudrun Socher, die meine Bachelorarbeit betreut und begutachtet hat. Für die hilfreichen Anregungen und die konstruktive Kritik bei der Erstellung dieser Arbeit möchte ich mich herzlich bedanken.

Ebenfalls möchte ich mich bei meinen Betreuer in der Firma Philipp Kapinos und meinem Kollegen Johannes Steinleitner bedanken, die mir mit viel Geduld, Interesse und Hilfsbereitschaft zur Seite standen. Bedanken möchte ich mich für die zahlreichen interessanten Debatten und Ideen, die maßgeblich dazu beigetragen haben, dass diese Bachelorarbeit in dieser Form vorliegt.

Abschließend möchte ich mich bei meinen Eltern und meinem Bruder bedanken, die mir mein Studium durch ihre Unterstützung ermöglicht haben und stets ein offenes Ohr für mich hatten.

Inhaltsverzeichnis

Zusammenfassung	i
Danksagungen	ii
Abbildungsverzeichnis	iii
Tabellenverzeichnis	iv
Abkürzungsverzeichnis	v
Ausgerichtet	v
1 Einleitung	1
1.1 Das Unternehmen Brunata-Metrona	1
1.2 Ziel der Arbeit	2
2 Grundlagen	3
2.1 Aktuelle Anwendung	3
2.2 Aktueller Stand	6
2.3 Analyse der Daten	7
2.3.1 Maschinelles Lernen	7
2.3.2 Data Mining und Text Mining	8
2.3.3 Bewertung	9
2.4 CRISP-DM: Prozessmodell für Analyseprojekte	10
2.4.1 Business Understanding	12
2.4.2 Data Understanding	12
2.4.3 Data Preparation	12
2.4.4 Modeling	13

INHALTSVERZEICHNIS

2.4.5	Evaluation	13
2.4.6	Deployment	14
3	Stand der Technik	15
3.1	Elasticsearch und Kibana	16
3.2	Verwendung	17
4	Konzept und Implementierung der Arbeit	20
4.1	Datenerfassung	21
4.2	Datenvorverarbeitung	21
4.3	Datenextraktion	23
4.4	Visualisierung	24
4.5	Datenverarbeitung	25
5	Fazit	26
5.1	Zusammenfassung der Arbeit	26
5.2	Zukünftige Arbeit	26
	Anhang 1: Einige Extras	28
	Anhang 2: Noch mehr Extras	29
	Literatur	30

Abbildungsverzeichnis

2.1	Vereinfachte Darstellung des Zusammenhangs der Mobilanwendungen der Brunata	4
2.2	Screenshot aus der Anwendung Gemo	5
2.3	Architekturkomponenten bei Brunata	6
2.4	Ausschnitt aus einer Logdatei	7
2.5	Prozessdiagramm veranschaulicht die Beziehung zwischen den verschiedenen Phasen des CRISP-DM-Prozesses. Illustration von Gökem Bas, basierend auf IBN SPSS Modeler CRISP-DM Guide	11
3.1	Prozessdiagramm veranschaulicht Methoden und Aufbau des Text Mining mittels Elasticsearch und Kibana	18
4.1	Beispieldatensatz der Speicherung von Geräte Logs auf dem FTP-Server	21
4.2	Ausschnitt aus dem Grok Debugger. Veranschaulicht das Grok-Pattern und die strukturierten Daten.	22
4.3	Beispieldatensatz der Abspeicherung von Geräte Logs auf dem FTP-Server	24

Tabellenverzeichnis

Abkürzungsverzeichnis

Ausgerichtet

GEMO	GE meinsame MO ntagelösung
KDD	K nowledge D iscovery in D ata
JSON	J ava S cript O bject N otation

1 | Einleitung

Angesichts dessen, dass die Leistung von mobilen Anwendungen im hohen Maße die Popularität und die von Nutzern wahrgenommene Dienstqualität bestimmt¹, erfreut sich der Markt für Performance Evaluation steigender Beliebtheit². So steuert beispielweise Google die Sichtbarkeit von mobilen Anwendungen im Playstore in Abhängigkeit der Ergebnisse der Performance Evaluation[2]. Somit besteht ein hohes Interesse darin die Performance zu optimieren, da nicht nur die Nutzerfreundlichkeit verbessert wird, sondern auch die Sichtbarkeit der App erhöht wird. Ein Unterfeld der Performance Evaluation beschäftigt sich mit der Verarbeitung von Logs. Sie werden benutzt, um Ereignisse zu dokumentieren oder um Entwicklern die Suche nach Fehlern zu erleichtern. Hierzu werden verschiedene verschiedenen Logebenen[1], Zeitstempeln oder Schemata verwendet, wodurch mannigfache Möglichkeiten bestehen, das Geloggte zu verarbeiten. Sind Logs nicht einheitlich nach demselben Schema dokumentiert, gibt es unterschiedliche Strukturen zu verarbeiten. Das heißt, dass eine Verarbeitung von Logs extrem schwierig ist, weil sich die Komplexität erhöht.

1.1 Das Unternehmen Brunata-Metrone

Die BRUNATA METRONA GmbH & Co. KG hat ihren Ursprung im Jahr 1958. Dass wie heute üblich jeder Haushalt nur seinen tatsächlichen Wärme- und Was-

¹vgl. Performance Evaluation of Mobile Agents: Issues and Approaches S.148[14] & Vgl. Computer Network 33 S.1-16[17]

²vgl. Gartner Report Magic Square

serververbrauch bezahlen muss, war damals Anreiz und Idee für die Firma BRUNATA Wärmemesser GmbH & Co. KG. Mittlerweile ist das Unternehmen unter dem Namen BRUNATA METRONA GmbH & Co. KG. Tätig. Zum Geschäftsfeld der Brunata gehört unter anderem die Installation von fernablesbaren Messgeräten. Dazu gehören Heizkostenverteiler und Wasserzähler. Diese dienen zur Ermittlung der Heiz- und Wasserkosten. Den Mess- und Montagepartnern wird die eigens entwickelte Mobile Anwendung Gemo zur Verfügung gestellt. Die Anwendung steuert dabei von der Planung eines Auftrags bis zur Ablesung einer Liegenschaft alle anfallenden Aufgaben eines Monteurs.

1.2 Ziel der Arbeit

Jedes Jahr erstellt die Brunata Abrechnungen für ca. 125.000 Liegenschaften. Nach einem Zeitraum von 12 Monaten werden die Verbrauchsstände dann von einem Ableser erfasst. Die gesammelten Daten werden mithilfe der eigens entwickelten Gemo-App verarbeitet und anschließend ins SAP-System übertragen. Da bei der Nutzung der App auch Fehler entstehen können, werden Abstürze der App automatisch gemeldet. Manche dieser Fehler führen nicht zu einem Absturz der App, sondern beeinträchtigen lediglich die Leistung wie z. B. eine lange Ladezeit. Diese Fehler müssten von Monteur selbstständig nach Abschluss der Arbeit gemeldet werden. Das hat zur Folge das Performanzprobleme lange unerkannt bleiben, Arbeitsprozesse länger brauchen als notwendig und erhöhter Personal und Kosten aufwand zur Behandlung und Ermittlung dieser Probleme notwendig ist. Das wirft die Problematik auf, dass nicht alle Performanzprobleme, die im Realbetrieb auftreten können, durch den aktuellen Umfang an Tests erkannt werden. Das Ziel dieser Arbeit ist es ein Verfahren zu Implementieren, um bisher unerkannte Leistungsprobleme wie beispielsweise lange Ladezeiten zu erkennen und zu beheben. Dazu werden aktuelle Verfahren und Techniken analysiert und eine passende Implementiert.

2 | Grundlagen

Die Performance Evaluation Mobiler Anwendungen ist der Prozess zur Bewertung der Qualität, Zuverlässigkeit und Leistung einer mobilen Anwendung, um ihre Eignung für einen bestimmten Zweck zu bestimmen. Dieser Prozess umfasst in der Regel das Sammeln und Analysieren von Metriken wie Reaktionszeit, Speichernutzung, Energieverbrauch und Benutzererfahrung sowie anderer Faktoren. Die Bewertung hilft Entwicklern, fundierte Entscheidungen über Design und die Funktionalität ihrer Produkte zu treffen, und kann auch dazu verwendet werden, verbesserungswürdige Bereiche zu ermitteln.[14]

2.1 Aktuelle Anwendung

Die Brunata nutzt zur Steuerung und Unterstützung der Arbeitsschritte der Mess- und Montagepartner eigens entwickelte Mobile Anwendungen. Die Mobile Anwendung Gemo bündelt dabei die Funktionen der anderen Anwendungen und ist die Schnittstelle für Mess- und Montage Partner. Der Aufbau und Zusammenhang der Anwendungen wird in Abbildung 2.1 Veranschaulicht.

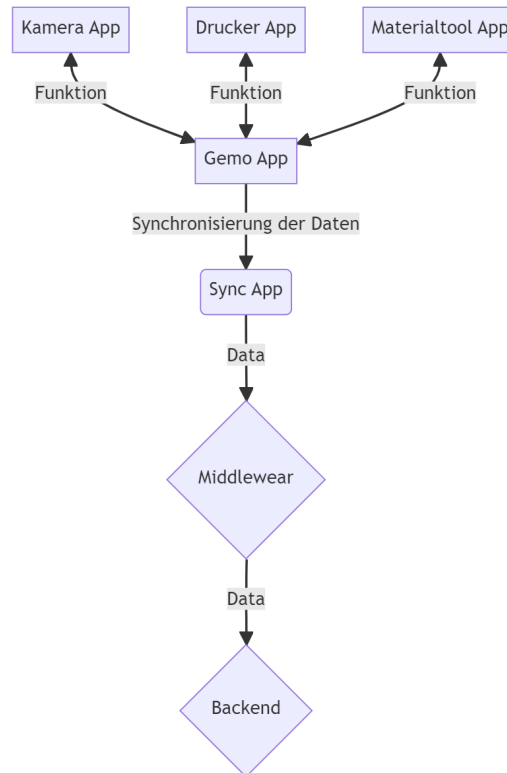


Abbildung 2.1: Vereinfachte Darstellung des Zusammenhangs der Mobilten Anwendungen der Brunata

Die Mobile Anwendung Gemo stellt unter anderem die Funktionen zur Verfügung Termine für die Ablesung von Zählerwerten sowie der Montage von Geräten zu erstellen. Des Weiteren leitet und kontrolliert die Anwendung den korrekten Ein- und Ausbau der Geräte und stellt sicher, dass Fehler erkannt und behoben werden können. Der Aufbau und die Funktionen der Anwendung sind in Abbildung 2.2 zu sehen.

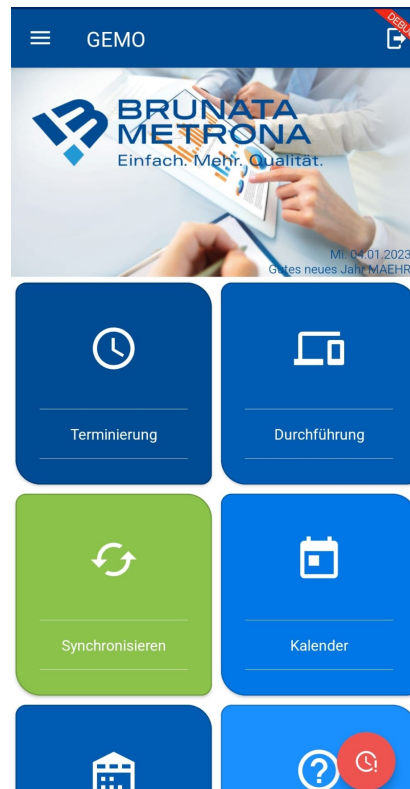


Abbildung 2.2: Screenshot aus der Anwendung Gemo

Die gesammelten Daten werden über die Sync-App Synchronisieren und ins SAP-System geladen. Zur geschützten Datenübertragung wird das mobile Internet mit der VPN-Technologie¹ verwendet. Wie in Abbildung 2.3 dargestellt, wird bei der Synchronisation von Daten über einen mobilen Client diese mittels RSA verschlüsselt und in ein text-basiertes Format (XML) in die Middleware übertragen. Die Anwendungsplattform SAP NetWeaver 7.x wird als Middlewarekomponente genutzt. Mit dieser Anwendung können unterschiedliche Geschäftsprozesse auf einer Plattform integriert werden.[15] Mithilfe der Middleware werden Backenddaten, welche speziell für die Benutzung mobiler Anwendungen aufbereitet wurden (SAP NetWeaver Mobile) verwaltet. Die Middleware Komponente dient vor allem der Verteilung von Benutzer- und Geräteverwaltung sowie Daten- und Softwareaktualisierungen. Die Daten aus der Middleware werden über einen XML-Parser für

¹VPN Virtual private Network ist ein geschlossenes Netzwerk, welches zum Transport privater Daten ein anderes öffentliches Netzwerk nutzt.[23]

die Anbindung an das ERP System (Backend) decodiert. Mithilfe von RFC² führt der Parser bestimmte Funktionen aus und strukturiert die Daten.

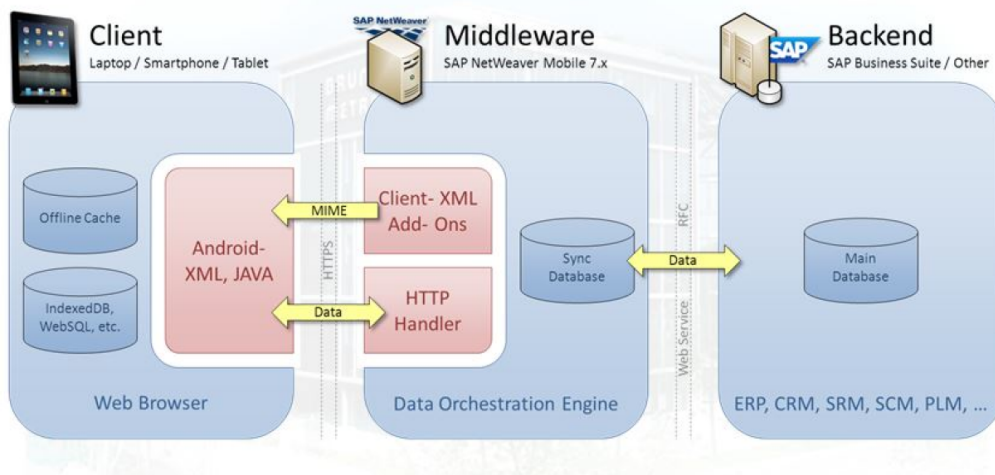


Abbildung 2.3: Architekturkomponenten bei Brunata

2.2 Aktueller Stand

Bei der Meldung eines Performanzproblems durch den Monteur werden die Logdaten und die Datenbank des Geräts auf den internen FTP-Server der Brunata geladen. Das gemeldete Problem wird von einem Entwickler überprüft. Dazu untersucht der Entwickler die hochgeladenen Logdaten des Geräts nach dem Problem. Dabei dient die benötigte Zeit von Prozessen und Anfragen, die geloggt wird als Anhaltspunkt für Leistungsprobleme. Die Logdateien umfassen ca. 5000 Zeilen und sind wie Abbildung 2.4 darstellt, nach dem Schema:

[Zeitstempel im ISO8601 Format] [Klasse oder Prozess des Aufrufs] [Log Level]
[Ereignis Beschreibung] generiert.

²Remote Function Call: SAP-Schnittstelle ermöglicht Kommunikation zwischen Systemen und führt Funktionen aus.[11]

```
2022-10-05 15:52:47.763 OmissionViewModel [DEBUG] Loading additional scheduleOrderData for B25CA78F43061EDCBDD786AC16648027
2022-10-05 15:52:47.764 HomeScreenFragment [DEBUG] updateFabVisibility()
2022-10-05 15:52:47.764 OmissionViewModel [DEBUG] loaded 69 items
2022-10-05 15:52:47.764 OmissionViewModel [INFO] countUnregisteredAppointments took: 349ms
2022-10-05 15:52:47.765 HomeScreenFragment [DEBUG] updateFabVisibility()
2022-10-05 15:52:47.766 UnfinishedTaskCount [DEBUG] countScheduleOrderStates
```

Abbildung 2.4: Ausschnitt aus einer Logdatei

Die Ereignisbeschreibung für die benötigte Zeit von Prozessen und Anfragen ist dabei nach dem Schema:

[Prozessname oder Anfrage] [“took”] [benötigte Zeit in ms] aufgebaut.

Durch die Größe der Logdaten und ihrer Auswertung durch einen Entwickler ist der aktuelle Prozess zur Erkennung und Behandlung von Leistungsproblemen zeit-aufwändig. Daher sollte die Untersuchung der Leistungsdaten vereinfacht werden. Dies kann durch die automatisierte Extraktion und Aufbereitung der Leistungsdaten aus den Logdaten realisiert werden.

2.3 Analyse der Daten

Es gibt eine Vielzahl von Analysetechniken, die zur Extraktion von Informationen aus Daten eingesetzt werden können. Zu diesen Techniken gehören maschinelles Lernen, Data Mining und Text Mining.

2.3.1 Maschinelles Lernen

Maschinelles Lernen ist ein Teilgebiet der künstlichen Intelligenz. Es befasst sich mit Algorithmen, welche auf Basis von Trainingsdaten automatisiert erwünschtes Verhalten erlernen.[16] Dabei sollte das Prinzip der Generalisierung gewahrt werde. Das bedeutet, ein erstelltes Modell kann auch auf unbekannte Daten das gewünschte Verhalten liefern.

Es wird beim maschinellen Lernen zwischen verschiedenen Arten des Lernens unterschieden:

1. Überwachtes Lernen
2. Unüberwachtes Lernen
3. Bestärkendes Lernen

Beim überwachten Lernen wird anhand von klassifizierten Daten ein Modell trainiert, das unbekannte Eingabewerte auf sinnvolle Zielwerte abbildet.[20] Beim unüberwachten Lernen dagegen werden die Daten nicht klassifiziert. Das Modell soll Strukturen, sogenannte Cluster, in den Daten erkennen. Es gibt dabei kein Zielwert als Vorgabe.[20] Beim bestärkenden Lernen wird das Modell durch ein Belohnungssystem trainiert. Dazu wird dem Modell durch Belohnung und Bestrafung mitgeteilt, ob seine Entscheidungen korrekt oder falsch waren. Dies hat zur Folge, dass zukünftige Entscheidungen entweder weiterverfolgt oder geändert werden[16]

Maschinelles Lernen kommt in unterschiedlichen Anwendungen zum Einsatz, in welchen Erkenntnisse aus Daten genutzt werden zur Lösung Praktischer Probleme.[20] So kann maschinelles Lernen dazu verwendet werden ein Modell zu Trainieren um Muster und Informationen aus Daten zur extrahieren und diese zu Bewerten. Allerdings ist die Modellerstellung aufwendig und die Güte des Modells hängt von der Qualität und Quantität der verfügbaren Daten ab.[20]

2.3.2 Data Mining und Text Mining

Data Mining (auch als engl. knowledge discovery in databases (KDD) bezeichnet)[20] ist ein Prozess, um Muster und Strukturen aus strukturierten Datenbeständen zu extrahieren.[12] Es gibt verschiedene Methoden für das Data Mining, dazu gehören Clusteranalysen, Bayes-Statistik sowie maschinelles Lernen.[12]

Die Clusteranalyse versucht, aufgrund der zwischen den Objekten vorhandenen Proximität eine gegebene Anzahl von Objekten in Gruppen aufzuteilen oder zu Gruppen zusammenzufassen, wobei sowohl eine herkömmliche als auch eine konzeptuelle Clustermethode angewendet werden kann.[12] Die Cluster werden so

konzipiert, dass die Ähnlichkeit der Objekte innerhalb einer Klasse maximiert und die Ähnlichkeit zwischen Objekten verschiedener Klassen minimiert wird.[12] Das Bayes-Verfahren beruht auf dem bekannten Theorem von Bayes über bedingte Wahrscheinlichkeiten, welches versucht, die Klassifikation so vorzunehmen, dass die bedingte Wahrscheinlichkeit für die Übereinstimmung der Klassenbildung mit der tatsächlichen Datenstruktur anhand der vorliegenden Daten maximiert wird. [12] Mithilfe von Maschinellern wird ein Modell trainiert, das ein Objekt aufgrund seiner Attribute in eine vorhandene Klasse einteilt. Techniken dazu wurden im Abschnitt 2.3.2 Maschinelles Lernen behandelt.

Text Mining ist ein Teilgebiet von Data Mining. Dabei werden keine strukturierten Daten, sondern unstrukturierte Texte verarbeitet. Bei der Verarbeitung von Text gibt es einige Besonderheiten im Vergleich zu strukturierten Daten. Text Mining hat dazu eine Reihe von spezifischen Methoden und Ansätze.[19] Die Methoden setzen sich dabei aus Kombinationen verwandter Verfahren, Bereiche wie Data Mining, maschinellem Lernen und natural language processing, zusammen. [19] Das Textmining wird verwendet zur Extraktion nicht trivialer Muster oder Inhalte aus unstrukturierten oder teilweise strukturierten Texten. Der Prozess für das Textmining ist ähnlich wie beim Data Mining. Sie unterscheiden sich in den Verfahren, die im Prozess verwendet werden.

Voraussetzung für ein aussagekräftiges Ergebnis bei Data und Text Mining ist die Qualität der Daten. Fehlerhafte oder unvollständige Daten wirken sich negativ auf das Ergebnis aus. Der Schutz der Daten vor Missbrauch und Diebstahl sowie die Einhaltung der geltenden Datenschutzverordnung bei der Verarbeitung sind beim Data Mining zu beachten und eine Herausforderung.

2.3.3 Bewertung

Der Aufwand und die Komplexität ein Modell auf Grundlage von Maschinellern zu trainieren ist im Vergleich zu den Alternativen groß. Obwohl das Verfahren maschinellem Lernen in der Lage wäre Informationen aus den Logdaten zu extrahieren und diese sogar automatisiert zu bewerten wird in Hinblick auf den

Aufwand, der Komplexität und der Unsicherheit der Qualität des Ergebnisses maschinelles Lernen nicht zur Lösung der Problemstellung verwendet. Allerdings könnte die Verwendung dieses Verfahrens in einer weiterführenden Arbeit untersucht werden. Aufgrund der Anforderung, dass Informationen, aus Logdaten, welche in Textform vorliegen, extrahiert werden sollen und Text Mining genau dazu verwendet wird, kommt im weiteren Verlauf der Arbeit Text Mining zum Einsatz.

2.4 CRISP-DM: Prozessmodell für Analyseprojekte

Ein standardisiertes und strukturiertes Vorgehen ermöglicht es, den Einstieg in die Durchführung von Analyseprojekten zu erleichtern und diese zu beschleunigen, die Planung und Verwaltung der erforderlichen Aktivitäten zu vereinfachen und die Validität und Verlässlichkeit der Ergebnisse zu erhöhen. Darüber hinaus fördert ein solches Prozessmodell die Wiederholbarkeit und Nachvollziehbarkeit aller ausgeführten Verarbeitungsschritte und Entscheidungen, was die Dokumentation und Wiederverwendung von Erfahrungen erleichtert. Schließlich hilft ein gemeinsames Verständnis über den allgemeinen Ablauf einer Datenanalyse bei der Kommunikation zwischen allen Beteiligten. Der Cross Industry Standard Process for Data Mining (CRISP-DM) ist ein sehr weit verbreitetes und ausgereiftes Prozessmodell, das ein Analyseprojekt hierarchisch auf vier Abstraktionsebenen herunterbricht. Auf der obersten Ebene wird der Analyseprozess, wie in Abbildung 2.5 dargestellt, in sechs Phasen unterteilt und abgedeckt.[21]

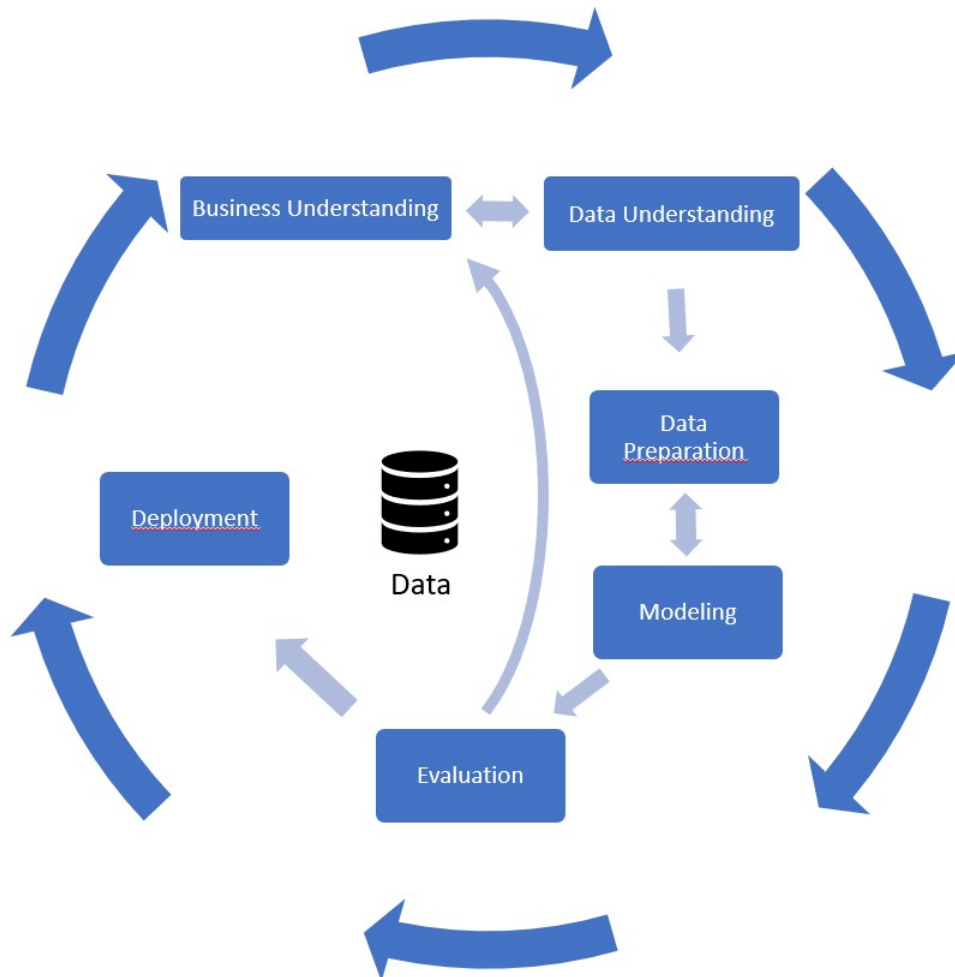


Abbildung 2.5: Prozessdiagramm veranschaulicht die Beziehung zwischen den verschiedenen Phasen des CRISP-DM-Prozesses. Illustration von Gökem Bas, basierend auf IBN SPSS Modeler CRISP-DM Guide

In dieser Arbeit wird ein Fokus auf die fachlichen und analytischen Komponenten der Phasen gelegt. Für eine umfassendere Erläuterung der Thematik ist auf die ausführliche CRISP-DM Dokumentation verwiesen [13]. Der Analyseprozess beginnt mit dem Business Understanding, folgt aber keinem streng linearen Ablauf. Während der Analyse neugewonnene Erkenntnisse erfordern häufig den mehrfachen Besuch einzelner Phasen. Wenn inhaltlich erforderlich ist ein beliebiger Sprung von einer Phase zur anderen möglich. Die Pfeile im Schaubild zeigen die in einem Analyseprojekt häufigsten zu beobachtenden Sprünge an. Der äußere

Kreislauf symbolisiert, dass der gesamte Analyseprozess oft mehrfach durchlaufen wird.[13] [21]

2.4.1 Business Understanding

Der Analyseprozess beginnt mit dieser Phase dabei ist es notwendig, den fachlichen Kontext und die Rahmenbedingungen des Anwendungsszenarios zu verstehen und die fachlichen Anforderungen und Ziele mit angemessenen Erfolgskriterien zu bestimmen. Es ist von besonderer Bedeutung zu verstehen, warum dieses Projekt für das Unternehmen von Nutzen ist sowie wie die Ergebnisse der Analyse verwendet werden sollen. Ohne messbare Zielsetzungen ist eine spätere Evaluierung der Analyseergebnisse problematisch. Erfahrungen aus vorhandenen Problembehandlungen innerhalb des Unternehmens sollten in Betracht gezogen werden.

2.4.2 Data Understanding

In der Data-Understanding-Phase erfolgt eine Initialanalyse der zur Verfügung stehenden Daten. Es wird eine Identifizierung relevanter und geeigneter Datenquellen vorgenommen und erste Erkenntnisse über die Daten gewonnen. Die mangelnde Datenqualität sowie die häufig begrenzten Möglichkeiten, diese zu verbessern, stellen ein signifikantes Risiko für eine erfolgreiche Analyse dar. Daher ist es erforderlich, dass Maßnahmen zur Erhöhung der Datenqualität getroffen werden. Dazu gehört auch, alternative Datenquellen zu erschließen, um die Relevanz, Menge und Qualität der Datenbasis zu erhöhen. Ziele und Erwartungshaltungen können durch einen Sprung in die erste Phase korrigiert werden.

2.4.3 Data Preparation

Die Data-Preparation-Phase hat das Ziel, Daten in einem Format aufzubereiten, welches für die nachfolgenden Methoden der Modellierungsphase angemessen ist.

Die Datenaufbereitung umfasst eine Reihe von Schritten, darunter die Integration von Daten aus verschiedenen Quellen, die Auswahl der zu verwendenden Datensätze und Merkmale, die Bereinigung, Transformationen, die Normalisierung und das Ableiten neuer Merkmale sowie die Formatierung. Während die Bereinigung und Transformation eine semantische Änderung der Daten beinhaltet, bezieht sich die Formatierung ausschließlich auf Änderungen syntaktischer Natur, die dazu dienen, die Daten für die vorgesehenen Methoden und Werkzeuge aufzubereiten. Die Datenaufbereitung gilt als Zentraler Erfolgsfaktor von Analyseprojekten. Da verschiedene Verarbeitungsmethoden unterschiedliche Anforderungen an die Daten stellen, gibt es keine Ideale Daten Aufbereitung. Diese wird durch ein Wechselspiel zwischen Datenvorbereitung und der Folgenden Modellierungsphase ermittelt.

2.4.4 Modeling

In der Modellierungsphase werden, geeignete Methoden für die jeweilige Analyseaufgabe ausgewählt und angewendet. Dabei müssen verschiedene Einflussfaktoren berücksichtigt werden, wie etwa die Beschaffenheit der Daten, die Verfügbarkeit von Methoden in entsprechenden Analyse-Werkzeugen und letztlich Anforderungen des Unternehmens. Wird eine Methode als adäquat erachtet, wird sie zum Zwecke der Evaluierung in die nächste Phase überführt.

2.4.5 Evaluation

Es wird allgemein angenommen, dass sich auch in rein zufällig erzeugten Datenbeständen durch intensive Suche Muster identifizieren lassen, die statistisch signifikant sind.[21] Die Evaluierung dieser Muster hinsichtlich ihrer Anwendbarkeit und Relevanz für ein Unternehmen ist jedoch unerlässlich, bevor sie zur Lösung einer fachlichen Fragestellung kommuniziert und angewendet werden. Der Zweck der Evaluierung ist es, den Nutzen der Analyseergebnisse im Kontext der Anwendung zu überprüfen und letztlich die Methode auszuwählen, die die Anforderungen unter Berücksichtigung der Erfolgskriterien am besten erfüllt. Dazu ist

es notwendig, dass die Ziele und Erfolgskriterien, die in der Phase des Business Understandings definiert wurden, berücksichtigt werden. Darüber hinaus sollte in dieser Phase der gesamte bisherige Analyseprozess einer Prüfung unterzogen werden. Die Evaluierungsmethoden und -kriterien hängen stark von der spezifischen Analyseaufgabe ab.

2.4.6 Deployment

Diese Phase stellt sicher, dass die Ergebnisse der Analyse den Anforderungen entsprechen, die mit der geplanten Anwendung verbunden sind. Der Begriff "Deployment" wird als Namensgeber für diese Phase verwendet und bezieht sich auf die Planung und Umsetzung der Anwendung der Ergebnisse. Die Spannbreite der möglichen Nutzung der Ergebnisse reicht von einem einmaligen Erkenntnisgewinn durch Interpretation und Verstehen über eine einmalige, manuelle Anwendung eines Modells auf einen ausgewählten Datenbestand bis hin zu einer regelmäßigen, automatischen Anwendung, die durch die Integration in betroffene Geschäftsprozesse realisiert wird. Bei einer regelmäßigen Anwendung ist es erforderlich, ein Modell zu kontrollieren und zu warten, um auf Veränderungen im Anwendungskontext oder der Umwelt angemessen reagieren zu können.[21]

3 | Stand der Technik

Performance relevante Daten zu erheben und verarbeiten ist eine bekannte Herausforderung bei der Performance Evaluation. Es gibt viele Lösungsansätze zu den verschiedenen Verfahren. Sie unterscheiden sich im Aufwand der Vorverarbeitung als auch in der Weiterverarbeitung der Daten. Die Schwierigkeit dabei ist es, die Leistung genau zu messen, da es eine Vielzahl von Geräten, Netzwerken und Betriebssystemen gibt, die verwendet werden.

Es hat sich in den letzten Jahren durchgesetzt, mittels Text Mining Leistungsdaten aus Logs zu extrahieren und grafisch aufzuarbeiten. Die Entwicklung einer solchen Auswertung und Analyse bei beispielweise Datadog[8], Dynatrace[9] und New Relic[10] hat zu einem Paradigmenwechsel in der Performance Evaluation geführt, sodass die Performance Evaluation für eine Vielzahl von Anwendungen ohne großen Aufwand möglich ist.

Bei Softwarelösungen der führenden Unternehmen wie Datadog ist es notwendig, zur vollumfänglichen Nutzung Daten außerhalb des Unternehmens zu verarbeiten.¹ Die Unternehmensphilosophie, das Daten nur im Unternehmen verarbeitet werden und die in Data Mining beschriebene Problematik in der datenschutzkonformen Datenverarbeitung führen dazu, dass keine Softwarelösungen mit externer Datenverarbeitung infrage kommen. Daher wird zur Implementierung des Text Mining Prozesses die Open-Source Such- und Analysesoftware Elasticsearch in Kombination mit der Open-Source-Analyse- und Visualisierungsplattform Kibana verwendet. Ziel der Implementierung ist es die benötigte Zeit von Prozessen

¹vgl. Magic Quadrant im Gartner Report [18] und Dokumentation Datadog [8]

und Anfragen aus den Logdaten zu Extrahieren und sie in form von Graphen darzustellen. Dadurch soll es Entwicklern ermöglicht werden bisher unerkannte Leistungsprobleme zu erkennen und beheben zu können.

3.1 Elasticsearch und Kibana

Elasticsearch ist eine verteilte Such- und Analysesoftware, die im Mittelpunkt des Elastic Stacks steht. Elasticsearch bietet Such- und Analysefunktionen für verschiedene Arten von Daten. Unabhängig davon, ob es sich um strukturierte oder unstrukturierte Text-, numerische oder Geodaten handelt, kann Elasticsearch diese Daten effizient speichern und indizieren, sodass eine schnelle Suche möglich ist. Elasticsearch bietet Methoden an, die über die einfache Datenabfrage hinausgehen und Informationen aggregieren, um Trends und Muster in den Daten erkennen. Die verteilte Natur von Elasticsearch ermöglicht es, dass sich der Einsatz an das wachsende Volumen von Daten und Abfragen anpasst.[7]

Beim Mapping werden die Felder der Daten, die indiziert werden sollen, definiert. Dieser Prozess ist notwendig, um sicherzustellen, dass die Suchmaschine die Daten richtig analysieren und interpretieren kann. Durch die Bereitstellung von Mapping-Informationen können Benutzer steuern, wie ihre Daten indiziert werden und wie die Suchmaschine die Daten bei einer Abfrage interpretiert.

Zur Vorbereitung und Strukturierung der Daten in die definierten Felder werden die Daten mit einem Grok Pattern welches auf den Inhalt der Daten passt extrahiert und Verarbeitet. Grok ist ein Dialekt für reguläre Ausdrücke, es verwendet die Oniguruma Bibliothek womit alle regulären Ausdrücke akzeptiert werden[4]. Grok erlaubt, bestehende Muster zu benennen und durch Kombination komplexere Muster zu erstellen, welche den gewollten Feldern entsprechen.

Auf den Indexierten Daten können zur Laufzeit Runtimefields erzeugt werden. Ein Runtimefield ist ein Feld das zur Laufzeit ausgewertet wird. Es dient dazu Felder ohne erneute Indexierung der Daten hinzuzufügen, den Rückgabewert eines Feldes zu überschreiben oder ein Feld für einen bestimmten Zweck zu

definieren ohne das zugrunde liegende Schema zu verändern[3].

Kibana ist eine Open-Source-Analyse- und Visualisierungsplattform für die Erstellung von interaktiven Dashboards und Grafiken auf Basis von Elasticsearch zur Erforschung und Visualisierung von Daten.

oder

Kibana ist eine Open-Source-Analyse- und Visualisierungsplattform, die es Benutzern ermöglicht, interaktive Dashboards und Grafiken zur Erforschung und Visualisierung von Daten auf Basis von Elasticsearch zu erstellen.

3.2 Verwendung

Mit einem kombinierten Einsatz der Methoden und Verfahren von Elasticsearch und Kibana ist es möglich, Leistungsdaten aus Logdaten zu extrahieren und grafisch aufzuarbeiten. Dadurch können Leistungsprobleme identifiziert und behoben werden. Die Abbildung 3.1 veranschaulicht, welche Aufgaben Elasticsearch und Kibana im Text Mining Prozess übernehmen.

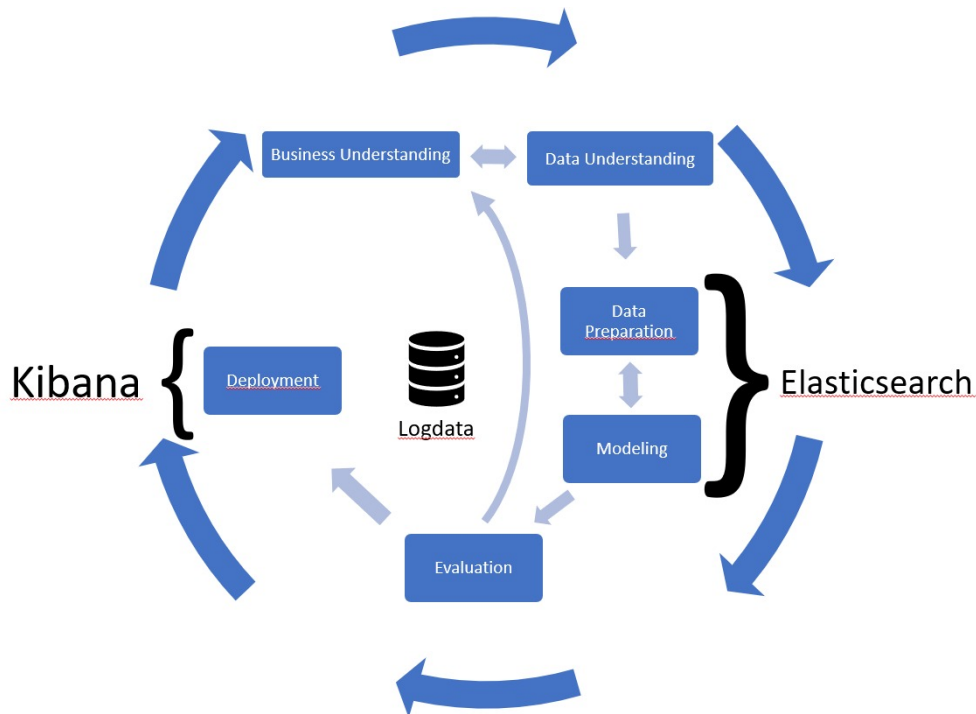


Abbildung 3.1: Prozessdiagramm veranschaulicht Methoden und Aufbau des Text Mining mittels Elasticsearch und Kibana

Elasticsearch übernimmt dabei die Vorbereitung der Daten und die Extraktion der Informationen. Kibana stellt die Visualisierung der Daten zur Verfügung. Nachdem Elasticsearch und Kibana Open-Source sind, können diese auf internen Server agieren und bereitgestellt werden.

<!--> Elasticsearch ist eine Opensource verteilte Such- und Analysemaschine für alle Arten von Daten. Unabhängig davon, ob es sich um strukturierten oder unstrukturierten Daten handelt, elasticsearch speicher die Daten Effizient und indiziert sie. Elasticsearch bietet neben der reinen Datenabfrage Methoden an um trends und Muster in den Daten zu erkenne. Elasticsearch basiert auf der Apache Lucene-Bibliothek, die einen invertierten Index verwendet, um Daten zu speichern und schnell darauf zuzugreifen. Elasticsearch ist hoch skalierbar und ermöglicht schnelle und effiziente Suchabfragen.

Bei der Entwicklung einer automatisierten First-Level-Support Schnittstelle, basierend auf maschinellen Lernen, hat die Firma Brunata-Metrona GmbH & Co

Kg gute Erfahrungen gemacht. Dabei ging es um die automatische Text Klassifikation von Support Anfragen bei der internen Support Hotline. Aufgrund dieser zufriedenstellenden Arbeit, soll im Rahmen dieser Arbeit erforscht werden, ob die Nutzung von Logs zur Performance Evaluation auch die gewünschten Ergebnisse erreicht.

Das Ziel dieser Arbeit ist es, basierend auf den Logs einer Mobilen Anwendung der Firma eine Performance Evaluation zu entwerfen und durchzuführen, die weiterentwickelt und auf weitere Anwendungen der Firma angewendet werden kann.

4 | Konzept und Implementierung der Arbeit

Um Leistungsprobleme mobiler Anwendungen erkennen und beheben zu können, soll die Leistung grafisch visualisiert und evaluiert werden. Dazu sollen Ladezeiten von Prozesse und Anfragen aus den Logdaten der Anwendung extrahiert werden. Zur Realisierung werden die Logdaten mithilfe der Open-Source Software Elasticsearch strukturiert, indexiert und anschließend die Leistungsdaten extrahiert. Für die Visualisierung werden die Daten mithilfe von Kibana in Form von Graphen aufgearbeitet und als Dashboard zur Verfügung gestellt. Im Rahmen dieser Arbeit werden für die Implementierung Logdaten der Anwendung Gemo verwendet.

4.1 Datenerfassung

Server: /clients/ABL0351T010_QSA/de.brunata.android.gemo/logs						
<ul style="list-style-type: none"> ABL0351T010_QR ABL0351T010_QSA <ul style="list-style-type: none"> de.brunata.android.cameraapp de.brunata.android.driver.metro de.brunata.android.ftp de.brunata.android.gemo <ul style="list-style-type: none"> backups logs de.brunata.android.geoha de.brunata.android.hkat de.brunata.android.mail de.brunata.android.materialtool de.brunata.android.metertool 						
Dateiname	Dateigröße	Dateityp	Zuletzt geändert	Berechtigu...	Besitzer/Gr...	
logs_2022-10-22_09-53-30_0.zip	10.793	WinRAR-ZI...	22.10.2022 09:5...	adfrw (0644)	1008 1008	
logs_2022-10-26_09-21-02_9.zip	11.795	WinRAR-ZI...	26.10.2022 09:2...	adfrw (0644)	1008 1008	
logs_2022-10-26_10-01-27_6.zip	426	WinRAR-ZI...	26.10.2022 10:0...	adfrw (0644)	1008 1008	
logs_2022-10-26_10-07-29_3.zip	11.214	WinRAR-ZI...	26.10.2022 10:0...	adfrw (0644)	1008 1008	

Abbildung 4.1: Beispieldatensatz der Speicherung von Geräte Logs auf dem FTP-Server

Bisher loggt die Anwendung die benötigte Zeit zum Abschließen von Performance relevanten Prozessen. Dazu wird beim Aufruf eines Prozesses eine Stoppuhr gestartet, welche mit Beendigung des Prozesses die verstrichene Zeit im Log dokumentiert. Intern ist festgelegt, dass am Ende des Tages und im Falle eines Geräte oder Anwendung Absturzes die Logs des Gerätes automatisch auf den internen FTP-Server der Brunata geladen werden. Jeder eingegangene Satz an Logdaten wird unter dem Gerätenamen sowie die Logs untergeordnet unter dem Anwendungsnamen abgespeichert. Ein Beispiel ist in Abbildung 4.3 zu sehen.

4.2 Datenvorverarbeitung

Zur Vorverarbeitung werden die Logdaten mit einem Gropattern in Felder strukturiert. Grok ist ein Dialekt für reguläre Ausdrücke, es verwendet die Oniguruma

Bibliothek womit alle regulären Ausdrücke akzeptiert werden[4]. Grok erlaubt, bestehende Muster zu benennen und durch Kombination komplexere Muster zu erstellen, welche den gewünschten Feldern entsprechen. Zur Strukturierung der Logdaten wird folgendes Grokpattern verwendet

```
1 %{TIMESTAMP_ISO8601:timestamp} %{DATA:class}
   .?*\[%{LOGLEVEL:loglevel}\] %{GREEDYDATA:message}
```

Das Pattern wird dabei auf jede Logzeile angewendet. Abbildung 4.2 veranschaulicht die Strukturierung durch das Grokpattern anhand einer Beispiel Auswertung im Grok Debugger.

The screenshot shows the Grok Debugger interface. At the top, under 'Sample Data', there is a table with one row: '2022-10-07 12:11:20.532 InstService [DEBUG] Loading of InstHeaders took 3ms'. Below this, under 'Grok Pattern', the pattern is entered: '%{TIMESTAMP_ISO8601:timestamp} %{DATA:class} .?*\[%{LOGLEVEL:loglevel}\] %{GREEDYDATA:message}'. A 'Custom Patterns' section is visible below. A blue 'Simulate' button is present. Under 'Structured Data', the result is shown as a JSON object: {'message': 'Loading of InstHeaders took 3ms', 'class': 'InstService', 'loglevel': 'DEBUG', 'timestamp': '2022-10-07 12:11:20.532'}.

Abbildung 4.2: Ausschnitt aus dem Grok Debugger. Veranschaulicht das Grokpattern und die Strukturierten Daten.

Ein Mapping bestimmt anschließend die Datentypen der Felder. Dazu wird die folgende Mappingvorlage verwendet:

```
1 "properties": {
   "@timestamp": { "type": "date" },
3  "log.level": { "type": "keyword" },
   "message": { "type": "wildcard" },
```

```
5 "class": { "type": "text" }
```

Listing 4.1: Mapping Vorlage der Daten

Die Auflistung 4.1 zeigt die Zuweisung der Datentypen zu, denn durch das Gropattern erstellten Feldern. Die Datentypen bestimmen die Eigenschaften der Felder und ihre interne Repräsentation. Der Datentyp “date” steht dabei für ein Zeitformat oder Datum Format. “keyword” wird für strukturierten Inhalt wie IDs, E-Mail Adressen oder Status Codes verwendet. “wildcard” ist für unstrukturierten Maschinen generierten Inhalt, er ist optimiert für Felder mit großen Werten oder hoher Kardinalität.[6] “text” ist für normalen Volltext Inhalt[5]. Der Datentyp bestimmt zusätzlich, ob und wie Felder weiterverarbeitet werden können.

4.3 Datenextraktion

Die benötigte Zeit von Prozessen und Anfragen sind nach der Vorverarbeitung im Feld “message” enthalten. Um die benötigte Zeit aus dem Feld zu extrahieren, werden Runtimefields erzeugt. Ein Runtimefield dient dazu, zur Laufzeit Ausdrücke aus Feldern zu extrahieren ohne erneute Indexierung[3]. Dazu werden Skripte, die Gropattern zur Auswertung des Feldes verwenden, angewendet. Das Runtimefield zur Extraktion der Benötigten zeit bekommt den Namen “duration” und den Datentyp “Double”. Zur vereinfachten Nutzung im weiteren Verlauf wird dem Runtimefield ein lable mit seinem Namen zugewiesen.

```
1 String duration=grok('%{WORD:typ} took
    %{NUMBER:duration}') .extract(doc["message"].value)?.duration;
    if (duration != null)
        emit(Integer.parseInt(duration));
```

Listing 4.2: Script zur Extraktion der benötigten Zeit

In der Auflistung 4.2 Zeile 1 wird ein Gropattern verwendet, um das Feld “message” nach dem in Aktueller Stand angesprochenen Schema auszuwerten und in Felder namens “typ” und “duration” zu strukturieren. Zusätzlich wird der wert

des Feldes “duration” extrahieren und in der Variable ‘duration’ gespeichert. Anschließend wird in Zeile 2 überprüft, ob die Variable ‘duration’ nicht null ist, und wenn das der Fall ist, wird durch Umwandlung der Variable der integer wert davon zurückgegeben. Die Abbildung ?? veranschaulicht den Erstellungsprozess in Elasticsearch.

4.4 Visualisierung

Die erstellten Felder werden in Kibana über ein Data View behandelt und verarbeitet. Über die grafische Oberfläche, die Kibana zur Verfügung stellt, werden Graphen auf Basis der Felder erstellt.

— Hier kommt bild von Dashboards rein Danach erklären wie die aufgebaut sind das wars—

<!-- # Konzept der Anwendung <!--

Datenerfassung

| Server: /clients/ABL0351T010_QSA/de.brunata.android.gemo/logs | | | | | | |
|---|-----------------------------------|--------------|--------------------|---------------|----------------|--|
| ? | ABL0351T010_QR | | | | | |
| ? | ABL0351T010_QSA | | | | | |
| ? | de.brunata.android.cameraapp | | | | | |
| ? | de.brunata.android.driver.metrona | | | | | |
| + | de.brunata.android.ftp | | | | | |
| + | de.brunata.android.gemo | | | | | |
| ? | backups | | | | | |
| | logs | | | | | |
| ? | de.brunata.android.geoha | | | | | |
| ? | de.brunata.android.hkat | | | | | |
| + | de.brunata.android.mail | | | | | |
| ? | de.brunata.android.materialtool | | | | | |
| ? | de.brunata.android.metertool | | | | | |
| | | | | | | |
| Dateiname | Dateigröße | Dateityp | Zuletzt geändert | Berechtigu... | Besitzer/Gr... | |
| .. | | | | | | |
| logs_2022-10-22-09-53-30_0.zip | 10.793 | WinRAR-ZI... | 22.10.2022 09:5... | adfrw (0644) | 1008 1008 | |
| logs_2022-10-26-09-21-02_9.zip | 11.795 | WinRAR-ZI... | 26.10.2022 09:2... | adfrw (0644) | 1008 1008 | |
| logs_2022-10-26-10-01-27_6.zip | 426 | WinRAR-ZI... | 26.10.2022 10:0... | adfrw (0644) | 1008 1008 | |
| logs_2022-10-26-10-07-29_3.zip | 11.214 | WinRAR-ZI... | 26.10.2022 10:0... | adfrw (0644) | 1008 1008 | |

Abbildung 4.3: Beispieldatensatz der Abspeicherung von Geräte Logs auf dem FTP-Server

Bisher loggt die Anwendung, die benötigte Zeit zum Abschließen von Performance Relevanten Prozessen. Dazu wird beim Aufruf eines Prozesses eine Stoppuhr gestartet, welche mit Beendigung des Prozesses die verstrichene Zeit im Log Dokumentiert. Intern ist festgelegt, dass am ende des Tages und im Falle eines Geräte oder Anwendung Absturzes die Logs des Gerätes automatisch auf den Internen FTP-Server der Brunata geladen werden. Jeder eingegangenen Satz an Log Daten wird unter dem Gerätenamen, sowie die Logs untergeordnet unter dem Anwendungsnamen, abgespeichert. Ein Beispiel ist in Abbildung 4.3 zu sehen. Im Rahmen dieser Arbeit werden nur Logs der Anwendung GEMO betrachtet.

4.5 Datenverarbeitung

Für die Verarbeitung der Logdaten wird Elasticsearch benutzt[22].

5 | Fazit

5.1 Zusammenfassung der Arbeit

Das Ziel dieser Bachelorarbeit war es, bisher unerkannte Leistungsprobleme in den Mobilanwendungen der Firma Brunata zu erkennen, um sie beheben zu können. Zur Realisierung wurde die Ausgangslage der Anwendungen untersucht, mögliche Verfahren zur Lösung der Problemstellung analysiert und ein Verfahren zur Erfüllung der Zielstellung implementiert.

Die Untersuchung zeigt, dass die von den Anwendungen dokumentierte benötigte Zeit von Prozessen und Anfragen als Indikator für Leistungsprobleme dient. Die Extraktion benötigter Zeit von Prozessen und Anfragen aus den Logdaten wird durch einen Text Mining Prozess realisiert. Zur Implementierung des Prozesses wird Elasticsearch und Kibana verwendet.

Zusammenfassend lässt sich festhalten, dass die Implementierung eines Text Mining Prozesses mithilfe von Elasticsearch und Kibana die Leistungsdaten von Anwendungen in grafischer Form darstellt und durch deren Auswertung Leistungsprobleme identifiziert.

5.2 Zukünftige Arbeit

Die hier dargestellten Ergebnisse führen zu der Frage, ist es möglich, ein Modell zu trainieren, das ein Leistungsproblem selbständig erkennt und meldet? Zur

Grafischen Auswertung wird ein Mitarbeiter benötigt welcher in dieser Zeit seiner normalen Tätigkeit nicht nachgehen kann. Würde nun ein Trainiertes Modell diese Aufgabe übernehmen wäre die Auswertung durch einen Mitarbeiter nicht mehr benötigt und es würden weiter Ressourcen gespart werden.

Die hier dargestellten Ergebnisse führen zu der Frage, wie würde man Leistungsdaten zur Erkennung von Leistungsproblemen ermitteln wenn die benötigte Zeit von Prozessen und Anfragen nicht Dokumentiert wird?

Für die Beantwortung dieser Frage ist eine detailliertere Untersuchung der Anwendung erforderlich um ein passendes Verfahren zu ermitteln und zu entwerfen das zum gewünschten Ziel führt.

Die hier dargestellten Ergebnisse führen zu der Frage was wäre, wenn die Leistungsdaten nicht durch Extraktion aus den logdaten, sondern mittels Schnittstellen aufgenommen und Verarbeitet werden? Die Daten werden nicht mehr einfach extrahiert und Grafisch aufbereitet sondern werden von einem Algorithmus bewertet. Die technischen Möglichkeiten, Daten aufgrund von ihren Eigenschaften zu bewerten, existieren bereits. Diese Technik wird Künstliche Inteligenz genannt und ist bereits im Intern

Für die Beantwortung, dieser Frage ist eine detailliertere Untersuchung erforderlich, die über einfache information's Extraktion und Darstellung hinausgeht. In einer solchen Studie könnte Durch die vorgestellten Ergebnisse lassen sich weiter Fragen aufwerfen: Was wäre wenn eine Auswertung nicht auf gemesGrafisch durch einen Entwickler, sondern automatisiert durch einen Algo

Anhang 1: Einige Extras

Füge Anhang 1 hier hinzu. Vivamus hendrerit rhoncus interdum. Sed ullamcorper et augue at porta. Suspendisse facilisis imperdiet urna, eu pellentesque purus suscipit in. Integer dignissim mattis ex aliquam blandit. Curabitur lobortis quam varius turpis ultrices egestas.

Anhang 2: Noch mehr Extras

Füge Anhang 2 hier hinzu. Aliquam rhoncus mauris ac neque imperdiet, in mattis eros aliquam. Etiam sed massa et risus posuere rutrum vel et mauris. Integer id mauris sed arcu venenatis finibus. Etiam nec hendrerit purus, sed cursus nunc. Pellentesque ac luctus magna. Aenean non posuere enim, nec hendrerit lacus. Etiam lacinia facilisis tempor. Aenean dictum nunc id felis rhoncus aliquam.

Literatur

- [1] Log ; ; android developers. *Android Developers*. Abgerufen von <https://developer.android.com/reference/android/util/Log#summary>
- [2] Android vitals. *Android Developers*. Abgerufen von <https://developer.android.com/topic/performance/vitals>
- [3] Explore your data with runtime fields. *Elastic Docs*. Abgerufen von <https://www.elastic.co/guide/en/elasticsearch/reference/current/runtime-examples.html>
- [4] How to write Scripts. *Elastic Docs*. Abgerufen von <https://www.elastic.co/guide/en/elasticsearch/reference/current/grok.html>
- [5] Text type family. *Elastic Docs*. Abgerufen von <https://www.elastic.co/guide/en/elasticsearch/reference/8.5/text.html>
- [6] Keyword type family. *Elastic Docs*. Abgerufen von <https://www.elastic.co/guide/en/elasticsearch/reference/8.5/keyword.html>
- [7] What is elasticsearch? *Elastic Docs*. Abgerufen von <https://www.elastic.co/guide/en/elasticsearch/reference/current/elasticsearch-intro.html>
- [8] Log Management. *Datadog Docs*. Abgerufen von <https://docs.datadoghq.com/logs/>
- [9] Log Management and Analytics. *Dynatrace Docs*. Abgerufen von <https://www.dynatrace.com/support/help/how-to-use-dynatrace/log-management-and-analytics>
- [10] Get started with log management. *Docs*. Abgerufen von <https://docs.newrelic.com/docs/logs/get-started/get-started-log-management/>
- [11] RFC. Abgerufen von https://help.sap.com/docs/SAP_NETWEAVER_700/109c9fd96c53101484f0ceb38844e91e/4888068ad9134076e1000000a42189d.html?locale=en-US&version=7.0.37&q=RFC
- [12] Hagedorn Jürgen Bissantz Nicolas. 2009. Data Mining (Datenmustererkennung). *WIRTSCHAFTSINFORMATIK 51*, 139–144. Abgerufen von <https://doi.org/10.1007/s11576-008-0108-z>

-
- [13] Peter Chapman, Janet Clinton, Randy Kerber, Tom Khabaza, Thomas P. Reinartz, Colin Shearer, und Richard Wirth. *CRISP-DM 1.0: Step-by-step data mining guide*.
 - [14] Marios D. Dikaiakos und George Samaras. 2001. Performance Evaluation of Mobile Agents: Issues and Approaches. In *Performance Engineering: State of the Art and Current Trends*, Reiner Dumke, Claus Rautenstrauch, André Scholz und Andreas Schmietendorf (Hrsg.). Springer Berlin Heidelberg, Berlin, Heidelberg, 148–166. Abgerufen 4. Dezember 2022 von https://doi.org/10.1007/3-540-45156-0_10
 - [15] Gisela Gratzl Erich Weirich Franz Josef Heiss. 2005. *SAP NetWeaver Web Application Server*. Addison-Wesley Verlag.
 - [16] Hartmut Ernst, Jochen Schmidt, und Gerd Beneken. 2020. Maschinelles Lernen: Deep Learning mit neuronalen Netzen. In *Grundkurs Informatik: Grundlagen und Konzepte für die erfolgreiche IT-Praxis – Eine umfassende, praxisorientierte Einführung*. Springer Fachmedien Wiesbaden, Wiesbaden, 801–833. DOI:https://doi.org/10.1007/978-3-658-30331-0_18
 - [17] Allan Kuchinsky Nina Bhatti Anna Bouch. 2000. *Integrating user-perceived quality into Web server design*. Elsevier Science B.V.
 - [18] Mrudula Bangera Padraig Byrne Gregg Siegfried. Magic Quadrant for Application Performance Monitoring and Observability. *Gartner*. Abgerufen von <https://www.gartner.com/doc/reprints?id=1-2A8RQQW4&ct=220608&st=sb>
 - [19] Claude Sammut und Geoffrey I. Webb (Hrsg.). 2017. Data mining on Text. In *Encyclopedia of Machine Learning and Data Mining*. Springer US, Boston, MA, 317–317. DOI:https://doi.org/10.1007/978-1-4899-7687-1_100099
 - [20] Carsten Lanquillon Sigurd Schacht (Hrsg.). 2019. *Blockchain und maschinelles Lernen*. Springer Vieweg Berlin, Heidelberg. DOI:<https://doi.org/https://doi-org.ezproxy.bib.hm.edu/10.1007/978-3-662-60408-3>
 - [21] Carsten Lanquillon Sigurd Schacht (Hrsg.). 2019. *Blockchain und maschinelles Lernen*. Springer Vieweg Berlin, Heidelberg. DOI:<https://doi.org/https://doi-org.ezproxy.bib.hm.edu/10.1007/978-3-662-60408-3>
 - [22] Armin Braun u.a. elasticsearch. Abgerufen von <https://github.com/elastic/elasticsearch>
 - [23] Zhang Zhang Zhensheng. 2004. An Overview of Virtual Private Network (VPN): IP VPN and Optical VPN. *Photonic Network Communications* 7, (2004), 213–225.
-