

Menschen Lernen Maschinelles Lernen

Image Text Matching

Seamless Analytics GmbH

Seamless Analytics GmbH

- Dienstleistungen und Beratung im Bereich der Datenspeicherung und –Analyse
- Unterstützung bei der Planung und Umsetzung von Projekten
 - Apache Hadoop Framework
 - Maschinelles Lernen
 - Gründung 2018
- Ansprechpartner
 - Marvin Follmann
 - Johannes Busam
 - Daniel Müller



Business Case

- Produktidee: Aufbau einer forensischen Plattform
 - Verteilte Speicherung und Parallelisierung über Cluster-Aufbau
 - Automatisierte Durchsuchung von Festplattenimages
 - Verwendung von ML zur Erweiterung der Analysemöglichkeiten
 - Verknüpfung von unstrukturierten Daten
 - Unterstützung bei forensischen Ermittlungen
 - Auffinden von Anomalien

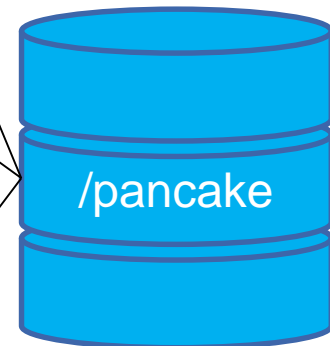
Aufgabenstellung

- Implementierung einer prototypischen Applikation, welche eine automatische Verknüpfung von Bild- und Textmaterial durchführt
 - Dient als PoC für spätere Analysen von digitalen Asservaten
 - Auffinden von Beziehungen zwischen Daten
(z.B. Verknüpfung von E-Mails mit Bildern bei forensischen Ermittlungen)
- Meilensteine
 - A: Automatisiertes Erstellen eines Trainingsdatensatzes mittels Verwendung von vortrainierten Neuronalen Netzen, um Relationsdaten von Text zu Bildern zu erhalten.
 - B: Trainieren eines Neuronalen Netzes welches vorhersagt, ob ein Textabschnitt und ein Bild zusammenpasst.
 - [C: Aufbau und Verwendung einer Taxonomie zur erweiterten Begriffsfindung (z.B. aus Wiktionary)]

- Als Beispieldatensatz dienen
 - Textuelle Produktbewertungen von Speisen
 - Bilder von Speisen
- Quelle: Kaggle.com

Gluten free pancakes

This flour is very nice in **pancakes**. I also used it in place of flour in making Apple Crisp, I will try other uses eventually.



Rahmenbedingungen

- Die nötigen Datensätze werden bereitgestellt.
- Lösungskonzept mit optionalen Ausbaustufen.
- Anwendung vortrainierter Neuronaler Netze.
- Das Projekt soll hauptsächlich an der Hochschule durchgeführt werden.
- Ein Team mit 2 Studenten wäre ideal.
- Als ML Framework soll TensorFlow dienen.

Wir freuen uns auf eure Unterstützung 😊

Daniel Müller

daniel.mueller@seamless-analytics.de

Marvin Follmann

marvin.follmann@seamless-analytics.de