

"Analyse der Byteverteilung in zlib-komprimierten Daten unter Variation der Fenstergröße mithilfe der Kullback-Leibler-Divergenz"

Diese Bachelorarbeit untersucht den Einfluss der Fenstergröße – ein Parameter im Deflate-Kompressionsverfahren, das zlib verwendet – auf die Byteverteilung der komprimierten Daten. Dabei steht insbesondere die Frage im Mittelpunkt, wie sich diese Byteverteilung bei Variation der Fenstergröße verändert und inwieweit sich solche Veränderungen quantitativ messen lassen. Gemeint sind hier mit Byteverteilung die relative Häufigkeiten der verschiedenen Bytewerte.

Zur Analyse wird die **Kullback-Leibler-Divergenz** eingesetzt, um Unterschiede zwischen den resultierenden Verteilungen zu quantifizieren. So lässt sich feststellen, ob bestimmte Fenstergrößen ähnliche oder stark unterschiedliche Verteilungen erzeugen. Zusätzlich zur KL-Divergenz zwischen den komprimierten Byteverteilungen soll auch die Divergenz zu einer Gleichverteilung betrachtet werden, um zu analysieren, wie zufällig eine komprimierte Datei ist (je nach Fenstergröße) und welche Fenstergröße eine Byteverteilung erzeugt, die näher an Entropie/Uniformität liegt. Experimente sollten für verschiedene Fenstergrößen und Eingabelängen durchgeführt werden.

Bei Interesse an dieser Bachelorarbeit bewerben Sie sich mit einem aussagekräftigen Exposé, das auch angibt, welcher Methodik Sie folgen wollen, bereits erfolgter Literaturrecherche und bestehendem (oder zumindest beginnendem) Grundlagenverständnis per Email.

Ihre Bewerbung sollte eine vorläufige Gliederung sowie einen stimmigen Zeitplan enthalten – zusätzlich zu Ihrem studentischen Werdegang und einer nachvollziehbaren Argumentation, warum Sie genau die richtige Person für dieses Thema sind.

Bewerbungen für diese Arbeit nehme ich bis Jahresende entgegen. Frühester Bearbeitungsbeginn soll Februar 2026 sein, später (spätestens Sommer 26) ist auch in Ordnung.