

---

# Übung zu der Vorlesung Text-Indexierung (WS20/21)

Hinweise

Abgabe bis spätestens Montag, den 08.02.2021 um 12:00 Uhr in elektronischer Form.

**Bitte geben Sie auch ein (leeres) PDF ab, wenn Sie nur die Programmieraufgabe lösen.** Dies hilft bei der Verwaltung der Punkte in Moodle.

Des Weiteren gelten alle organisatorischen Hinweise vom ersten Übungsblatt.

## Programmieraufgabe

(5 Punkte)

Die genaue Aufgabenstellung entnehmen Sie bitte der `README.md` des Aufgaben-Repositories. Um das Ihr Repository zu erzeugen folgen Sie bitte den folgenden Link (Link nur in der Moodle-internen Version).

- (a) Implementieren Sie ein Programm, welches als Eingabe einen Text  $T$  und einen Parameter  $k$  erhält und die  $k$ -te Entropie von  $T$  ausgibt. Hierbei sollte das Programm mindestens für  $k \leq 3$  funktionieren. (4 Punkte)

## Huffman und Shannon-Fano

(5 Punkte)

Zeigen Sie, dass (trotz der Optimalität) manche Codewörter bei Huffman ein längeres Codewort bekommen als bei Shannon-Fano. Warum widerspricht das nicht der Optimalität von Huffman?