

# NLLR

## Toteutusdokumentti

Leo Leppänen

17. helmikuuta 2014

### 1 Algoritmit

#### 1.1 Argmax.single

##### 1.1.1 Tilavaativuus

Tilavaativuus on  $\mathcal{O}(1)$ , sillä algoritmit käyttää vakiomäärää muuttujia ja tulostaa aina yhden Result-olion.

##### 1.1.2 Aikavaativuus

Argmax suorittaa syötteenään saamansa algoritmin kerran per syöteenä saadun argumenttilistan argumentti, joten aikavaativuus on  $\mathcal{O}(A \times n)$ , missä  $n$  on maksimoitavan argumentin kandidaattien määrä ja  $A$  on suoritettavan algoritmin aikavaativuus..

#### 1.2 Argmax.multiple

##### 1.2.1 Tilavaativuus

Tilavaativuus on  $\mathcal{O}(n)$ , sillä kerrallaan muistissa pidetään korkeintaan  $n + 1$  Result-oliota sekä vakiomäärää muita muuttujia.

### 1.2.2 Aikavaativuus

Algoritmi suorittaa syötteenä saadun  $A$  aikavaativuuksisen algoritmin  $n$  kertaa, jolloin tältä osin aikavaativuus on  $\mathcal{O}(A \times n)$ . Lisäksi pahimmillaan  $n$  kertaa kutsutaan metodia `sort()`, joka järjestää tuloslistan.

Järjestysalgoritmina toimii InsertionSort. Järjestysalgoritmin valintaan vaikutti uniikki konteksti: jokaisella järjestyskerralla kaikki paitsi yksi alkio ovat valmiina oikeilla paikoillaan. Lisäksi järjestettävät taulukot erittäin pienikokoisia. Näissä tapauksissa InsertionSort on nopein ja tehokkain <sup>1</sup>. Tässä erityistapauksessamme aikavaativuus on lähempänä  $\mathcal{O}(n)$  kuin  $\mathcal{O}(n^2)$  ja tilavaativuus on  $\mathcal{O}(1)$ .

## 1.3 NLLR

### 1.3.1 Tilavaativuus

$\mathcal{O}(1)$ , sillä algoritmi käyttää syötteensä lisäksi vain vakiomäärän tilaa `bestTokens`-taulukon sekä välitulokset tallentavien muuttujien muodossa.

### 1.3.2 Aikavaativuus

Algoritmi määrittää aluksi `Argmax`:lla vakiomäärän parhaan TF-IDF arvon saavia sanoja, joille sen jälkeen kullekin suoritetaan useita  $\mathcal{O}(1)$  aikavaativuuksisia `calculateTokenPropability`-komentoja. Täten aikavaativuus on sama kuin

---

<sup>1</sup><http://dl.acm.org/citation.cfm?doid=359024.359026>

## **1.4 TFIDF**

### **1.4.1 Tilavaativuus**

### **1.4.2 Aikavaativuus**

## **2 Tietorakenteet**

### **2.1 ArrayList**

#### **2.1.1 Tilavaativuus**

#### **2.1.2 Aikavaativuus**

### **2.2 HashMap**

#### **2.2.1 Tilavaativuus**

#### **2.2.2 Aikavaativuus**

### **2.3 HashSet**

#### **2.3.1 Tilavaativuus**

#### **2.3.2 Aikavaativuus**