



Maschinelle Sprachverarbeitung

Übung

Aufgabe 5: Evaluationsergebnisse

Mario Sängner

mario.saenger@informatik.hu-berlin.de

Übung 5

- Implementation eines regelbasierte Verfahrens zur Erkennung von Gen- / Proteinnamen in Texten
 - Fokus auf Single-Token Gennamen
 - Bereitstellung eines Wörterbuchs mit Genenamen + Trainingsdaten
- Verwendung von vielgestaltigen Eigenschaften
 - (Fuzzy-) Dictionary-Matching, Reguläre Ausdrücke, Stoppwörter, Wörterbucherweiterungen, Kontextwörter, ...
- Wettbewerb: Die Lösung mit dem höchsten F1-Score gewinnt!

Lösungsübersicht

Gruppe	Ansatz	
	Sprache	Features
AS_JG	Python	POS-Tags, N-Gram-Jaccard-Distanz
Batorurowaiaru	Python	Regel-Engine, Wortaufbau, Wörterbuch
Bitdiggers	Python	Wörterbuch, Stoppwörter
DieBesteGruppe	Python	Wörterbuch, Wortaufbau
NotWorkingAsIntended	Python	Erw. Wörterbuch, Stoppwörter
Pi	Python	Wortaufbau, Gewichtete Regeln
TraubeMinze&EinTee	Java	Stoppwörter, Bereinigtes Wörterbuch
WWGDS	Java	Wörterbuch

Trainingsergebnisse

Gruppe	P	R	F1	Rang
AS-JG	10.92	42.65	17.38	8
Batorurowaiaru	48.73	87.77	62.67	5
Bitdiggers	39.93	44.36	42.03	7
DieBesteGruppe	41.89	67.11	51.58	6
NotWorkingAsIntended	93.22	95.82	94.50	1
Pi	70.92	66.16	68.46	4
TraubeMinze&EinTee	57.06	99.11	72.42	2
WWGDS	54.90	99.43	70.74	3

Testergebnisse

Gruppe	P	R	F1
AS-JG	10.61	44.20	17.11
Batorurowaiaru	45.10	82.60	58.34
Bitdiggers	41.53	48.07	44.56
DieBesteGruppe	41.26	67.13	51.10
NotWorkingAsIntended	58.82	60.77	59.78
Pi	66.67	59.12	62.66
TraubeMinze&EinTee	44.89	65.47	53.26
WWGDS	43.35	65.75	52.25

Testergebnisse

Gruppe	P	R	F1
AS-JG	10.61	44.20	17.11
Batorurowaiaru	45.10	82.60	58.34
Bitdiggers	41.53	48.07	44.56
DieBesteGruppe	41.26	67.13	51.10
NotWorkingAsIntended	58.82	60.77	59.78
Pi	66.67	59.12	62.66
TraubeMinze&EinTee	44.89	65.47	53.26
WWGDS	43.35	65.75	52.25

Testergebnisse

Gruppe	P	R	F1	Rang	Punkte
AS-JG	10.61	44.20	17.11	8	
Batorurowaiaru	45.10	82.60	58.34	3	1
Bitdiggers	41.53	48.07	44.56	7	
DieBesteGruppe	41.26	67.13	51.10	6	
NotWorkingAsIntended	58.82	60.77	59.78	2	3
Pi	66.67	59.12	62.66	1	5
TraubeMinze&EinTee	44.89	65.47	53.26	4	
WWGDS	43.35	65.75	52.25	5	

Testergebnisse

Gruppe	TP	FP	FN	Rang
AS-JG	160	1348	202	8
Batorurowaiaru	299	364	63	3
Bitdiggers	174	245	188	7
DieBesteGruppe	243	346	119	6
NotWorkingAsIntended	220	154	142	2
Pi	214	107	148	1
TraubeMinze&EinTee	237	291	125	4
WWGDS	238	311	124	5

Road-Map für die letzten 2 Wochen

- Do, 07.02.2019
 - Optionaler Q&A-Termin
- So, 10.02.2019
 - Abgabe 6. Übungsaufgabe
- Do, 14.02.2019
 - Besprechung 6. Übungsaufgabe
 - Abschluss, Feedback, Siegerehrung