

# Knowledge-base for Word Prediction

---

Elior Matan : 308470715

Ibrahim Awad : 318260585

Statistics:

MapReduce Step		With local aggregation	Without local aggregation
Step 1 Divide	Mapper in	163471963	163471963
	Mapper out	163471963	163471963
	Combiner in	163471963	0
	Combiner out	3295187	0
	Reducer in	3295187	163471963
	Reducer out	3172966	3172966
Step 2 CalcProb	Mapper in	3172966	3172966
	Mapper out	6345929	6345929
	Combiner in	6345929	0
	Combiner out	14125	0
	Reducer in	14125	6345929
	Reducer out	9985	9985
Step 3 JoinResults	Mapper in	3182951	3182951
	Mapper out	3182951	3182951
	Combiner in	0	0
	Combiner out	0	0
	Reducer in	3182951	3182951
	Reducer out	3172966	3172966

Step 4 Sort	Mapper in	3172966	3172966
	Mapper out	3172966	3172966
	Combiner in	0	0
	Combiner out	0	0
	Reducer in	3172966	3172966
	Reducer out	3172966	3172966

Analysis:

1. הבדל זה

0.00000026846804361941    הבדל זה הוא  
0.00000018851444592948    הבדל זה אינו  
0.00000017311054178738    הבדל זה שבין  
0.00000015477256066583    הבדל זה .  
0.0000000990250980563    הבדל זה נובע

2. הבולט ביותר

0.00000045404841256944    הבולט ביותר הוא  
0.00000018264629197058    הבולט ביותר היה  
0.00000018044573423600    הבולט ביותר בין  
0.00000012983290634053    הבולט ביותר .  
0.00000004474467393657    הבולט ביותר לכך

3. הבונה את

0.00000004181059695712    הבונה את ארצו  
0.00000003814300073281    הבונה את עצמה  
0.00000003740948148795    הבונה את הבית  
0.00000003594244299823    הבונה את העולם  
0.00000003374188526364    הבונה את ביתו

4. יכול היה

0.00000163794847377634    יכול היה לראות  
0.00000156826414551447    יכול היה לעשות  
0.00000156606358777988    יכול היה לעמוד  
0.00000116776263781994    יכול היה לדעת  
0.00000103866325072427    יכול היה שלא

5. כתוצאה מכך

0.00000050686179819948 כתוצאה מכך .  
0.00000042617468126469 כתוצאה מכך לא  
0.00000021858873496881 כתוצאה מכך נוצר  
0.00000020098427309213 כתוצאה מכך הוא  
0.00000018191277272572 כתוצאה מכך היו

6. כתיב כי  
0.00000033081717943266 כתיב כי לא  
0.00000012689882936109 כתיב כי גדול  
0.00000011076140597413 כתיב כי אם  
0.00000009022286711800 כתיב כי הוא  
0.00000006308265505811 כתיב כי כל

7. כתיבתו של  
0.00000027727027455775 כתיבתו של ספר  
0.00000017824517650141 כתיבתו של הספר  
0.00000006234913581325 כתיבתו של עגנון  
0.00000005134634714032 כתיבתו של השיר  
0.00000004987930865060 כתיבתו של הסופר

8. לא -  
0.00000848608414380584 לא - יהודים  
0.00000303310207750343 לא - יהודי  
0.00000252257268307963 לא - מעט  
0.00000251303693289643 לא - אחת  
0.00000216754936856654 לא - פעם

9. על המשבר  
0.00000026846804361941 על המשבר .  
0.00000019071500366406 על המשבר הכלכלי  
0.00000009975861730120 על המשבר ומתה  
0.00000005868153958894 על המשבר של  
0.00000005207986638519 על המשבר הרוחני

10. שלא נוח  
0.00000018264629197058 שלא נוח לו  
0.00000007481896297590 שלא נוח היה  
0.00000005868153958894 שלא נוח להם  
0.00000004547819318143 שלא נוח לה  
0.00000004547819318143 שלא נוח לי