

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	96.71 %	64.29 %	63.64 %	98.51 %	11	202
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	9	178
C (إسم معرف)	93.84 %	95.81 %	33.33 %	97.81 %	9	137
D (إضافة)	92.81 %	68.67 %	85.71 %	93.6 %	14	125
E (منعوت)	97.53 %	96.92 %	66.67 %	98.72 %	6	156
F (نعت)	79.44 %	51.74 %	71.43 %	80.12 %	14	166
G (حرف جر)	93.18 %	48.24 %	0.0 %	95.35 %	1	43
H (إسم مجرور)	69.23 %	48.96 %	57.89 %	80.0 %	19	20
I (وحدة قياس)	98.96 %	49.74 %	100.0 %	98.94 %	2	94
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	26
K (فعل مبني للمجهول)	88.89 %	47.06 %	66.67 %	93.33 %	3	15
L (المفعول المطلق)	50.0 %	33.33 %	50.0 %	NULL	2	0
M (أداة عطف غير واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	1
.	100.0 %	100.0 %	NULL	100.0 %	0	392

Table 1: Performances for each tag using a first order HMM for fold 0

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	95.5 %	63.38 %	73.68 %	97.54 %	19	203
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	10	167
C (إسم معرف)	91.38 %	93.69 %	56.25 %	94.94 %	16	158
D (إضافة)	86.98 %	65.18 %	73.68 %	88.67 %	19	150
E (منعوت)	97.63 %	98.0 %	66.67 %	98.77 %	6	163
F (نعت)	77.25 %	64.62 %	65.22 %	78.92 %	23	166
G (حرف جر)	93.48 %	48.31 %	NULL	93.48 %	0	46
H (إسم مجرور)	65.0 %	49.24 %	35.29 %	86.96 %	17	23
I (وحدة قياس)	98.86 %	49.71 %	100.0 %	98.85 %	1	87
J (واو العطف)	97.06 %	49.25 %	NULL	97.06 %	0	34
K (فعل مبني للمجهول)	76.92 %	28.99 %	50.0 %	88.89 %	8	18
L (المفعول المطلق)	33.33 %	16.67 %	33.33 %	NULL	3	0
M (أداة عطف غير واو العطف)	80.0 %	44.44 %	NULL	80.0 %	0	5
.	100.0 %	100.0 %	NULL	100.0 %	0	406

Table 2: Performances for each tag using a first order HMM for fold 1

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	94.19 %	61.37 %	80.65 %	96.19 %	31	210
B (رقم)	98.68 %	74.66 %	80.0 %	99.32 %	5	146
C (إسم معرف)	92.59 %	92.16 %	66.67 %	95.83 %	18	144
D (إضافة)	89.44 %	67.2 %	76.47 %	90.97 %	17	144
E (منعوت)	99.33 %	99.45 %	75.0 %	100.0 %	4	146
F (نعت)	78.92 %	68.94 %	68.42 %	80.27 %	19	147
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	26
H (إسم مجرور)	78.95 %	53.57 %	40.0 %	92.86 %	5	14
I (وحدة قياس)	100.0 %	100.0 %	NULL	100.0 %	0	79
J (واو العطف)	96.3 %	49.06 %	NULL	96.3 %	0	27
K (فعل مبني للمجهول)	50.0 %	22.22 %	0.0 %	66.67 %	3	9
L (المفعول المطلق)	0.0 %	0.0 %	0.0 %	NULL	1	0
M (أداة عطف غير واو العطف)	83.33 %	45.45 %	NULL	83.33 %	0	6
.	100.0 %	100.0 %	NULL	100.0 %	0	406

Table 3: Performances for each tag using a first order HMM for fold 2

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	93.91 %	61.43 %	73.68 %	95.73 %	19	211
B (رقم)	98.4 %	49.6 %	80.0 %	98.9 %	5	182
C (إسم معرف)	90.73 %	61.28 %	60.0 %	94.12 %	15	136
D (إضافة)	89.33 %	88.76 %	46.15 %	93.43 %	13	137
E (منعوت)	97.84 %	65.16 %	75.0 %	99.24 %	8	131
F (نعت)	78.44 %	51.66 %	57.14 %	80.39 %	14	153
G (حرف جر)	96.55 %	49.12 %	50.0 %	100.0 %	2	27
H (إسم مجرور)	78.57 %	56.75 %	0.0 %	88.0 %	3	25
I (وحدة قياس)	98.04 %	33.0 %	100.0 %	98.0 %	2	100
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	18
K (فعل مبني للمجهول)	87.5 %	46.67 %	100.0 %	85.71 %	2	14
L (المفعول المطلق)	100.0 %	100.0 %	NULL	100.0 %	0	1
M (أداة عطف غير واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	8
.	100.0 %	100.0 %	NULL	100.0 %	0	393

Table 4: Performances for each tag using a first order HMM for fold 3

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	93.81 %	62.67 %	63.64 %	95.35 %	11	215
B (رقم)	99.52 %	49.88 %	100.0 %	99.49 %	12	196
C (إسم معرف)	90.17 %	61.62 %	68.75 %	92.36 %	16	157
D (إضافة)	89.53 %	88.49 %	64.29 %	91.77 %	14	158
E (منعوت)	97.1 %	64.49 %	70.0 %	99.22 %	10	128
F (نعت)	74.86 %	50.78 %	60.0 %	76.77 %	20	155
G (حرف جر)	94.74 %	32.43 %	50.0 %	97.22 %	2	36
H (إسم مجرور)	70.0 %	48.81 %	41.67 %	82.14 %	12	28
I (وحدة قياس)	98.02 %	55.22 %	100.0 %	97.94 %	4	97
J (واو العطف)	96.77 %	49.18 %	NULL	96.77 %	0	31
K (فعل مبني للمجهول)	84.0 %	45.65 %	100.0 %	82.61 %	2	23
L (المفعول المطلق)	100.0 %	100.0 %	NULL	100.0 %	0	1
M (أداة عطف غير واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	13
.	100.0 %	100.0 %	NULL	100.0 %	0	393

Table 5: Performances for each tag using a first order HMM for fold 4

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	96.4 %	63.63 %	81.82 %	97.16 %	11	211
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	6	205
C (إسم معرف)	96.18 %	97.43 %	80.0 %	97.89 %	15	142
D (إضافة)	92.36 %	67.93 %	85.71 %	93.01 %	14	143
E (منعوت)	100.0 %	100.0 %	100.0 %	100.0 %	5	141
F (نعت)	80.84 %	54.39 %	61.11 %	83.22 %	18	149
G (حرف جر)	97.87 %	49.46 %	NULL	97.87 %	0	47
H (إسم مجرور)	77.78 %	77.71 %	57.14 %	90.91 %	14	22
I (وحدة قياس)	100.0 %	100.0 %	100.0 %	100.0 %	3	114
J (واو العطف)	97.67 %	49.41 %	NULL	97.67 %	0	43
K (فعل مبني للمجهول)	79.17 %	58.19 %	66.67 %	80.95 %	3	21
L (المفعول المطلق)	100.0 %	100.0 %	NULL	100.0 %	0	1
M (أداة عطف غير واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	10
.	100.0 %	100.0 %	NULL	100.0 %	0	392

Table 6: Performances for each tag using a first order HMM for fold 5

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	97.22 %	96.1 %	61.54 %	99.51 %	13	203
B (رقم)	99.52 %	49.88 %	50.0 %	100.0 %	2	206
C (إسم معرف)	90.6 %	93.69 %	50.0 %	94.81 %	14	135
D (إضافة)	85.43 %	63.59 %	43.48 %	92.97 %	23	128
E (منعوت)	98.64 %	98.38 %	80.0 %	99.3 %	5	142
F (نعت)	79.75 %	53.39 %	60.0 %	81.08 %	10	148
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	51
H (إسم مجرور)	83.78 %	56.08 %	54.55 %	96.15 %	11	26
I (وحدة قياس)	100.0 %	100.0 %	100.0 %	100.0 %	2	111
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	37
K (فعل مبني للمجهول)	71.43 %	48.57 %	50.0 %	75.0 %	6	36
L (المفعول المطلق)	100.0 %	100.0 %	NULL	100.0 %	0	1
M (أداة عطف غير واو العطف)	85.71 %	46.15 %	NULL	85.71 %	0	7
.	100.0 %	100.0 %	NULL	100.0 %	0	392

Table 7: Performances for each tag using a first order HMM for fold 6

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	98.17 %	97.79 %	76.92 %	99.51 %	13	206
B (رقم)	99.48 %	49.87 %	0.0 %	100.0 %	1	190
C (إسم معرف)	92.17 %	61.52 %	75.0 %	94.0 %	16	150
D (إضافة)	89.88 %	89.89 %	56.0 %	95.8 %	25	143
E (منعوت)	96.55 %	64.6 %	60.0 %	99.26 %	10	135
F (نعت)	83.44 %	55.84 %	38.46 %	87.33 %	13	150
G (حرف جر)	100.0 %	100.0 %	100.0 %	100.0 %	1	46
H (إسم مجرور)	87.18 %	56.91 %	70.0 %	93.1 %	10	29
I (وحدة قياس)	100.0 %	100.0 %	100.0 %	100.0 %	1	73
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	31
K (فعل مبني للمجهول)	80.0 %	29.63 %	50.0 %	86.49 %	8	37
L (المفعول المطلق)	NULL	NULL	NULL	NULL	0	0
M (أداة عطف غير واو العطف)	75.0 %	42.86 %	NULL	75.0 %	0	8
.	100.0 %	100.0 %	NULL	100.0 %	0	392

Table 8: Performances for each tag using a first order HMM for fold 7



tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	98.16 %	65.66 %	91.67 %	98.54 %	12	205
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	3	190
C (إسم معرف)	94.67 %	63.46 %	81.25 %	96.08 %	16	153
D (إضافة)	94.08 %	93.38 %	82.35 %	95.39 %	17	152
E (منعوت)	97.35 %	65.25 %	33.33 %	98.65 %	3	148
F (نعت)	90.24 %	60.16 %	63.64 %	92.16 %	11	153
G (حرف جر)	100.0 %	100.0 %	100.0 %	100.0 %	1	59
H (إسم مجرور)	89.36 %	58.44 %	72.73 %	94.44 %	11	36
I (وحدة قياس)	100.0 %	100.0 %	100.0 %	100.0 %	1	79
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	41
K (فعل مبني للمجهول)	88.37 %	31.28 %	88.89 %	88.24 %	9	34
L (المفعول المطلق)	NULL	NULL	NULL	NULL	0	0
M (أداة عطف غير واو العطف)	83.33 %	45.45 %	NULL	83.33 %	0	12
.	100.0 %	100.0 %	NULL	100.0 %	0	390

Table 9: Performances for each tag using a first order HMM for fold 8

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	96.96 %	64.33 %	73.33 %	98.6 %	15	215
B (رقم)	99.47 %	49.87 %	50.0 %	100.0 %	2	186
C (إسم معرف)	91.41 %	91.41 %	66.67 %	94.48 %	18	145
D (إضافة)	85.45 %	84.79 %	47.37 %	90.41 %	19	146
E (منعوت)	97.69 %	97.55 %	66.67 %	99.19 %	6	124
F (نعت)	86.13 %	57.42 %	61.54 %	88.71 %	13	124
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	59
H (إسم مجرور)	83.33 %	54.87 %	71.43 %	88.24 %	14	34
I (وحدة قياس)	100.0 %	100.0 %	100.0 %	100.0 %	2	92
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	40
K (فعل مبني للمجهول)	84.48 %	45.79 %	75.0 %	86.0 %	8	50
L (المفعول المطلق)	NULL	NULL	NULL	NULL	0	0
M (أداة عطف غير واو العطف)	83.33 %	45.45 %	NULL	83.33 %	0	12
.	100.0 %	100.0 %	NULL	100.0 %	0	390

Table 10: Performances for each tag using a first order HMM for fold 9

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	96.1 %	70.07 %	74.06 %	97.66 %	16	208
B (رقم)	99.51 %	72.38 %	76.0 %	99.77 %	6	185
C (إسم معرف)	92.37 %	81.21 %	63.79 %	95.23 %	15	146
D (إضافة)	89.53 %	77.79 %	66.12 %	92.6 %	18	143
E (منعوت)	97.97 %	84.98 %	69.33 %	99.23 %	6	141
F (نعت)	80.93 %	56.89 %	60.7 %	82.9 %	16	151
G (حرف جر)	97.58 %	72.76 %	60.0 %	98.39 %	1	44
H (إسم مجرور)	78.32 %	56.13 %	50.07 %	89.28 %	12	26
I (وحدة قياس)	99.39 %	78.77 %	100.0 %	99.37 %	2	93
J (واو العطف)	98.78 %	79.69 %	NULL	98.78 %	0	33
K (فعل مبني للمجهول)	79.08 %	40.4 %	64.72 %	83.39 %	5	26
L (المفعول المطلق)	69.05 %	64.29 %	27.78 %	100.0 %	1	0
M (أداة عطف غير واو العطف)	89.07 %	66.98 %	NULL	89.07 %	0	8
.	100.0 %	100.0 %	NULL	100.0 %	0	395

Table 11: average performances for each tag using a first order HMM

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	unknown words percentage
0	94.77 %	67.67 %	67.78 %	96.33 %	78	9.67 %
1	93.02 %	65.56 %	63.93 %	95.2 %	107	13.26 %
2	94.15 %	67.44 %	69.9 %	95.81 %	88	10.9 %
3	94.01 %	87.89 %	62.65 %	95.7 %	75	9.29 %
4	93.14 %	87.59 %	67.96 %	94.73 %	92	11.4 %
5	95.78 %	69.61 %	76.4 %	96.83 %	83	10.29 %
6	94.38 %	67.53 %	54.65 %	96.49 %	80	9.91 %
7	95.08 %	91.17 %	61.22 %	97.17 %	90	11.15 %
8	96.83 %	94.04 %	79.76 %	97.7 %	77	9.54 %
9	94.92 %	90.7 %	64.95 %	96.72 %	88	10.9 %
Avg	94.61 %	78.92 %	66.92 %	96.27 %	86.0	10.63 %

Table 12: Average performances of first order HMM used to predict ingredients state using 10 fold cross-validation