

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	number of known words
A (إسم)	96.24 %	63.91 %	72.73 %	97.52 %	11	202
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	9	178
C (إسم معرف)	90.41 %	60.28 %	66.67 %	91.97 %	9	137
D (إضافة)	86.33 %	64.08 %	78.57 %	87.2 %	14	125
E (منعوت)	96.3 %	95.5 %	83.33 %	96.79 %	6	156
F (نعت)	64.44 %	42.92 %	50.0 %	65.66 %	14	166
G (حرف جر)	95.45 %	48.84 %	0.0 %	97.67 %	1	43
H (إسم مجرور)	66.67 %	47.44 %	52.63 %	80.0 %	19	20
I (وحدة قياس)	98.96 %	49.74 %	100.0 %	98.94 %	2	94
J (واو العطف)	96.15 %	49.02 %	NULL	96.15 %	0	26
K (فعل مبني للمجهول)	100.0 %	100.0 %	100.0 %	100.0 %	3	15
L (المفعول المطلق)	100.0 %	100.0 %	100.0 %	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	784

Table 1: Performances for each tag using a second 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)	96.24 %	63.91 %	72.73 %	97.52 %	11	202
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	9	178
C (إسم معرف)	90.41 %	60.28 %	66.67 %	91.97 %	9	137
D (إضافة)	86.33 %	64.08 %	78.57 %	87.2 %	14	125
E (منعوت)	96.3 %	95.5 %	83.33 %	96.79 %	6	156
F (نعت)	64.44 %	42.92 %	50.0 %	65.66 %	14	166
G (حرف جر)	95.45 %	48.84 %	0.0 %	97.67 %	1	43
H (إسم مجرور)	66.67 %	47.44 %	52.63 %	80.0 %	19	20
I (وحدة قياس)	98.96 %	49.74 %	100.0 %	98.94 %	2	94
J (واو العطف)	96.15 %	49.02 %	NULL	96.15 %	0	26
K (فعل مبني للمجهول)	100.0 %	100.0 %	100.0 %	100.0 %	3	15
L (المفعول المطلق)	100.0 %	100.0 %	100.0 %	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	784

Table 2: Performances for each tag using asecond 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 %	91.75 %	83.87 %	95.71 %	31	210
B (رقم)	99.34 %	83.17 %	100.0 %	99.32 %	5	146
C (إسم معرف)	91.36 %	90.45 %	83.33 %	92.36 %	18	144
D (إضافة)	83.85 %	62.98 %	58.82 %	86.81 %	17	144
E (منعوت)	97.33 %	97.85 %	75.0 %	97.95 %	4	146
F (نعت)	71.69 %	56.37 %	52.63 %	74.15 %	19	147
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	26
H (إسم مجرور)	68.42 %	46.23 %	40.0 %	78.57 %	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)	75.0 %	28.57 %	66.67 %	77.78 %	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	812

Table 3: Performances for each tag using a second 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)	94.78 %	62.28 %	94.74 %	94.79 %	19	211
B (رقم)	99.47 %	49.87 %	100.0 %	99.45 %	5	182
C (إسم معرف)	93.38 %	92.28 %	60.0 %	97.06 %	15	136
D (إضافة)	87.33 %	86.67 %	53.85 %	90.51 %	13	137
E (منعوت)	90.65 %	88.13 %	37.5 %	93.89 %	8	131
F (نعت)	68.26 %	47.29 %	50.0 %	69.93 %	14	153
G (حرف جر)	100.0 %	100.0 %	100.0 %	100.0 %	2	27
H (إسم مجرور)	71.43 %	49.44 %	66.67 %	72.0 %	3	25
I (وحدة قياس)	98.04 %	49.5 %	100.0 %	98.0 %	2	100
J (واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	18
K (فعل مبني للمجهول)	81.25 %	29.89 %	100.0 %	78.57 %	2	14
L (المفعول المطلق)	0.0 %	0.0 %	NULL	0.0 %	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	786

Table 4: Performances for each tag using a second 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)	92.04 %	60.62 %	54.55 %	93.95 %	11	215
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	12	196
C (إسم معرف)	93.64 %	92.96 %	68.75 %	96.18 %	16	157
D (إضافة)	87.21 %	85.69 %	71.43 %	88.61 %	14	158
E (منعوت)	92.75 %	90.6 %	60.0 %	95.31 %	10	128
F (نعت)	66.29 %	45.98 %	40.0 %	69.68 %	20	155
G (حرف جر)	97.37 %	49.33 %	100.0 %	97.22 %	2	36
H (إسم مجرور)	65.0 %	44.3 %	58.33 %	67.86 %	12	28
I (وحدة قياس)	99.01 %	83.08 %	100.0 %	98.97 %	4	97
J (واو العطف)	96.77 %	49.18 %	NULL	96.77 %	0	31
K (فعل مبني للمجهول)	92.0 %	47.92 %	100.0 %	91.3 %	2	23
L (المفعول المطلق)	0.0 %	0.0 %	NULL	0.0 %	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	786

Table 5: Performances for each tag using a second 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)	93.69 %	61.05 %	72.73 %	94.79 %	11	211
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	6	205
C (إسم معرف)	90.45 %	60.41 %	73.33 %	92.25 %	15	142
D (إضافة)	85.35 %	62.58 %	85.71 %	85.31 %	14	143
E (منعوت)	95.89 %	95.06 %	40.0 %	97.87 %	5	141
F (نعت)	68.26 %	47.36 %	27.78 %	73.15 %	18	149
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	47
H (إسم مجرور)	66.67 %	66.56 %	50.0 %	77.27 %	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)	75.0 %	39.44 %	66.67 %	76.19 %	3	21
L (المفعول المطلق)	0.0 %	0.0 %	NULL	0.0 %	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	784

Table 6: Performances for each tag using a second 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)	97.69 %	96.77 %	84.62 %	98.52 %	13	203
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	2	206
C (إسم معرف)	84.56 %	56.43 %	50.0 %	88.15 %	14	135
D (إضافة)	81.46 %	61.5 %	47.83 %	87.5 %	23	128
E (منعوت)	97.28 %	96.87 %	100.0 %	97.18 %	5	142
F (نعت)	75.32 %	51.7 %	60.0 %	76.35 %	10	148
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	51
H (إسم مجرور)	72.97 %	46.66 %	45.45 %	84.62 %	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)	50.0 %	25.78 %	50.0 %	50.0 %	6	36
L (المفعول المطلق)	0.0 %	0.0 %	NULL	0.0 %	0	1
M (أداة عطف غير واو العطف)	100.0 %	100.0 %	NULL	100.0 %	0	7
.	100.0 %	100.0 %	NULL	100.0 %	0	784

Table 7: Performances for each tag using asecond 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)	96.35 %	64.01 %	76.92 %	97.57 %	13	206
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	1	190
C (إسم معرف)	87.35 %	86.44 %	50.0 %	91.33 %	16	150
D (إضافة)	83.33 %	82.9 %	44.0 %	90.21 %	25	143
E (منعوت)	97.24 %	97.03 %	80.0 %	98.52 %	10	135
F (نعت)	80.37 %	55.91 %	38.46 %	84.0 %	13	150
G (حرف جر)	100.0 %	100.0 %	100.0 %	100.0 %	1	46
H (إسم مجرور)	74.36 %	46.06 %	50.0 %	82.76 %	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)	55.56 %	23.81 %	50.0 %	56.76 %	8	37
L (المفعول المطلق)	NULL	NULL	NULL	NULL	0	0
M (أداة عطف غير واو العطف)	87.5 %	46.67 %	NULL	87.5 %	0	8
.	100.0 %	100.0 %	NULL	100.0 %	0	784

Table 8: Performances for each tag using a second 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)	95.85 %	63.45 %	75.0 %	97.07 %	12	205
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	3	190
C (إسم معرف)	91.12 %	90.41 %	62.5 %	94.12 %	16	153
D (إضافة)	86.98 %	86.36 %	58.82 %	90.13 %	17	152
E (منعوت)	97.35 %	97.22 %	66.67 %	97.97 %	3	148
F (نعت)	84.76 %	58.35 %	54.55 %	86.93 %	11	153
G (حرف جر)	100.0 %	100.0 %	100.0 %	100.0 %	1	59
H (إسم مجرور)	70.21 %	44.18 %	27.27 %	83.33 %	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)	79.07 %	29.44 %	66.67 %	82.35 %	9	34
L (المفعول المطلق)	NULL	NULL	NULL	NULL	0	0
M (أداة عطف غير واو العطف)	91.67 %	47.83 %	NULL	91.67 %	0	12
.	100.0 %	100.0 %	NULL	100.0 %	0	780

Table 9: Performances for each tag using a second 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 %	80.0 %	98.14 %	15	215
B (رقم)	100.0 %	100.0 %	100.0 %	100.0 %	2	186
C (إسم معرف)	85.89 %	85.72 %	61.11 %	88.97 %	18	145
D (إضافة)	81.82 %	82.28 %	47.37 %	86.3 %	19	146
E (منعوت)	98.46 %	98.39 %	83.33 %	99.19 %	6	124
F (نعت)	85.4 %	57.73 %	76.92 %	86.29 %	13	124
G (حرف جر)	100.0 %	100.0 %	NULL	100.0 %	0	59
H (إسم مجرور)	68.75 %	43.61 %	35.71 %	82.35 %	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)	63.79 %	25.96 %	50.0 %	66.0 %	8	50
L (المفعول المطلق)	NULL	NULL	NULL	NULL	0	0
M (أداة عطف غير واو العطف)	91.67 %	47.83 %	NULL	91.67 %	0	12
.	100.0 %	100.0 %	NULL	100.0 %	0	780

Table 10: Performances for each tag using a second 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)	95.15 %	69.02 %	75.83 %	96.46 %	16	208
B (رقم)	99.88 %	93.3 %	100.0 %	99.88 %	6	185
C (إسم معرف)	89.72 %	77.41 %	65.07 %	92.29 %	15	146
D (إضافة)	84.59 %	73.64 %	60.43 %	87.79 %	18	143
E (منعوت)	95.97 %	95.37 %	70.92 %	97.16 %	6	141
F (نعت)	72.77 %	51.24 %	50.25 %	75.06 %	16	151
G (حرف جر)	98.85 %	84.71 %	80.0 %	99.05 %	1	44
H (إسم مجرور)	68.7 %	48.06 %	46.72 %	78.7 %	12	26
I (وحدة قياس)	99.49 %	83.2 %	100.0 %	99.48 %	2	93
J (واو العطف)	98.4 %	74.59 %	NULL	98.4 %	0	33
K (فعل مبني للمجهول)	76.78 %	39.98 %	73.75 %	77.9 %	5	26
L (المفعول المطلق)	23.81 %	20.0 %	55.56 %	0.0 %	1	0
M (أداة عطف غير واو العطف)	93.42 %	73.22 %	NULL	93.42 %	0	8
.	100.0 %	100.0 %	NULL	100.0 %	0	789

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

tag	accuracy	f1 score	accuracy for unknown words	accuracy for known words	number of unknown words	unknown words percentage
0	93.72 %	64.13 %	70.0 %	94.81 %	78	9.67 %
1	92.53 %	63.0 %	64.75 %	94.19 %	107	13.26 %
2	94.14 %	65.6 %	70.87 %	95.39 %	88	10.9 %
3	93.94 %	86.92 %	68.67 %	95.02 %	75	9.29 %
4	93.42 %	86.61 %	66.02 %	94.81 %	92	11.4 %
5	93.83 %	65.85 %	62.92 %	95.18 %	83	10.29 %
6	93.76 %	65.63 %	60.47 %	95.19 %	80	9.91 %
7	94.04 %	88.56 %	55.1 %	95.96 %	90	11.15 %
8	95.34 %	90.99 %	60.71 %	96.77 %	77	9.54 %
9	94.34 %	88.56 %	61.86 %	95.91 %	88	10.9 %
Avg	93.91 %	76.58 %	64.14 %	95.32 %	86.0	10.63 %

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