

| tag                              | accuracy | f1 score | accuracy for<br>unknown<br>words | accuracy for<br>known<br>words | number of<br>unknown<br>words | number of<br>known<br>words |
|----------------------------------|----------|----------|----------------------------------|--------------------------------|-------------------------------|-----------------------------|
| A<br>(إسم)                       | 83.33 %  | 12.99 %  | 63.64 %                          | 84.36 %                        | 11                            | 211                         |
| B<br>(رقم)                       | 97.63 %  | 19.76 %  | 83.33 %                          | 98.05 %                        | 6                             | 205                         |
| C<br>(إسم معرف)                  | 82.17 %  | 11.28 %  | 73.33 %                          | 83.1 %                         | 15                            | 142                         |
| D<br>(إضافة)                     | 82.8 %   | 12.94 %  | 64.29 %                          | 84.62 %                        | 14                            | 143                         |
| E<br>(منعوت)                     | 90.41 %  | 13.57 %  | 80.0 %                           | 90.78 %                        | 5                             | 141                         |
| F<br>(نعت)                       | 85.63 %  | 11.53 %  | 72.22 %                          | 87.25 %                        | 18                            | 149                         |
| G<br>(حرف جر)                    | 78.72 %  | 17.62 %  | NULL                             | 78.72 %                        | 0                             | 47                          |
| H<br>(إسم مجرور)                 | 72.22 %  | 16.77 %  | 42.86 %                          | 90.91 %                        | 14                            | 22                          |
| I<br>(وحدة قياس)                 | 93.16 %  | 24.12 %  | 33.33 %                          | 94.74 %                        | 3                             | 114                         |
| J<br>(واو العطف)                 | 90.7 %   | 19.02 %  | NULL                             | 90.7 %                         | 0                             | 43                          |
| K<br>(فعل مبني<br>للمجهول)       | 54.17 %  | 14.05 %  | 0.0 %                            | 61.9 %                         | 3                             | 21                          |
| L<br>(المفعول المطلق)            | 0.0 %    | 0.0 %    | NULL                             | 0.0 %                          | 0                             | 1                           |
| M<br>(أداة عطف غير<br>واو العطف) | 30.0 %   | 9.23 %   | NULL                             | 30.0 %                         | 0                             | 10                          |
| .                                | 100.0 %  | 100.0 %  | NULL                             | 100.0 %                        | 0                             | 784                         |

Table 1: Performances for each tag of our ingredient extractor for fold 5 and for the first layer with  $\lambda = 4$  using a first order HMM

| tag                              | accuracy | f1 score | accuracy for<br>unknown<br>words | accuracy for<br>known<br>words | number of<br>unknown<br>words | number of<br>known<br>words |
|----------------------------------|----------|----------|----------------------------------|--------------------------------|-------------------------------|-----------------------------|
| A<br>(إسم)                       | 94.59 %  | 62.35 %  | 81.82 %                          | 95.26 %                        | 11                            | 211                         |
| B<br>(رقم)                       | 98.58 %  | 49.64 %  | 100.0 %                          | 98.54 %                        | 6                             | 205                         |
| C<br>(إسم معرف)                  | 93.63 %  | 47.12 %  | 93.33 %                          | 93.66 %                        | 15                            | 142                         |
| D<br>(إضافة)                     | 84.71 %  | 59.96 %  | 42.86 %                          | 88.81 %                        | 14                            | 143                         |
| E<br>(منعوت)                     | 95.89 %  | 63.66 %  | 40.0 %                           | 97.87 %                        | 5                             | 141                         |
| F<br>(نعت)                       | 83.83 %  | 57.08 %  | 72.22 %                          | 85.23 %                        | 18                            | 149                         |
| G<br>(حرف جر)                    | 93.62 %  | 32.23 %  | NULL                             | 93.62 %                        | 0                             | 47                          |
| H<br>(إسم مجرور)                 | 66.67 %  | 44.99 %  | 50.0 %                           | 77.27 %                        | 14                            | 22                          |
| I<br>(وحدة قياس)                 | 100.0 %  | 100.0 %  | 100.0 %                          | 100.0 %                        | 3                             | 114                         |
| J<br>(واو العطف)                 | 95.35 %  | 48.81 %  | NULL                             | 95.35 %                        | 0                             | 43                          |
| K<br>(فعل مبني<br>للمجهول)       | 79.17 %  | 22.09 %  | 66.67 %                          | 80.95 %                        | 3                             | 21                          |
| L<br>(المفعول المطلق)            | 100.0 %  | 100.0 %  | NULL                             | 100.0 %                        | 0                             | 1                           |
| M<br>(أداة عطف غير<br>واو العطف) | 80.0 %   | 29.63 %  | NULL                             | 80.0 %                         | 0                             | 10                          |
| .                                | 99.49 %  | 33.25 %  | NULL                             | 99.49 %                        | 0                             | 784                         |

Table 2: Performances for each tag of our ingredient extractor for fold 5 and for the second layer with  $\lambda = 4$

|                                  | First Layer<br>first order<br>HMM |          |                              |                            | Second Layer<br>second order<br>HMM |          |                              |                            |
|----------------------------------|-----------------------------------|----------|------------------------------|----------------------------|-------------------------------------|----------|------------------------------|----------------------------|
| Tag                              | accuracy                          | f1 score | accuracy<br>unknown<br>words | accuracy<br>known<br>words | accuracy                            | f1 score | accuracy<br>unknown<br>words | accuracy<br>known<br>words |
| A<br>(إسم)                       | 83.33 %                           | 12.99 %  | 63.64 %                      | 84.36 %                    | 94.59 %                             | 62.35 %  | 81.82 %                      | 95.26 %                    |
| B<br>(رقم)                       | 97.63 %                           | 19.76 %  | 83.33 %                      | 98.05 %                    | 98.58 %                             | 49.64 %  | 100.0 %                      | 98.54 %                    |
| C<br>(إسم معرف)                  | 82.17 %                           | 11.28 %  | 73.33 %                      | 83.1 %                     | 93.63 %                             | 47.12 %  | 93.33 %                      | 93.66 %                    |
| D<br>(إضافة)                     | 82.8 %                            | 12.94 %  | 64.29 %                      | 84.62 %                    | 84.71 %                             | 59.96 %  | 42.86 %                      | 88.81 %                    |
| E<br>(منعوت)                     | 90.41 %                           | 13.57 %  | 80.0 %                       | 90.78 %                    | 95.89 %                             | 63.66 %  | 40.0 %                       | 97.87 %                    |
| F<br>(نعت)                       | 85.63 %                           | 11.53 %  | 72.22 %                      | 87.25 %                    | 83.83 %                             | 57.08 %  | 72.22 %                      | 85.23 %                    |
| G<br>(حرف جر)                    | 78.72 %                           | 17.62 %  | NULL                         | 78.72 %                    | 93.62 %                             | 32.23 %  | NULL                         | 93.62 %                    |
| H<br>(إسم مجرور)                 | 72.22 %                           | 16.77 %  | 42.86 %                      | 90.91 %                    | 66.67 %                             | 44.99 %  | 50.0 %                       | 77.27 %                    |
| I<br>(وحدة قياس)                 | 93.16 %                           | 24.12 %  | 33.33 %                      | 94.74 %                    | 100.0 %                             | 100.0 %  | 100.0 %                      | 100.0 %                    |
| J<br>(واو العطف)                 | 90.7 %                            | 19.02 %  | NULL                         | 90.7 %                     | 95.35 %                             | 48.81 %  | NULL                         | 95.35 %                    |
| K<br>(فعل مبني<br>للمجهول)       | 54.17 %                           | 14.05 %  | 0.0 %                        | 61.9 %                     | 79.17 %                             | 22.09 %  | 66.67 %                      | 80.95 %                    |
| L<br>(المفعول المطلق)            | 0.0 %                             | 0.0 %    | NULL                         | 0.0 %                      | 100.0 %                             | 100.0 %  | NULL                         | 100.0 %                    |
| M<br>(أداة عطف غير<br>واو العطف) | 30.0 %                            | 9.23 %   | NULL                         | 30.0 %                     | 80.0 %                              | 29.63 %  | NULL                         | 80.0 %                     |
| .                                | 100.0 %                           | 100.0 %  | NULL                         | 100.0 %                    | 99.49 %                             | 33.25 %  | NULL                         | 99.49 %                    |

Table 3: Performances for each tag of our ingredient extractor for fold 5 and for the first and second layer with  $\lambda = 4$

|      | First Layer<br>first order<br>HMM |          |                              |                            | Second Layer<br>second order<br>HMM |          |                              |                            |
|------|-----------------------------------|----------|------------------------------|----------------------------|-------------------------------------|----------|------------------------------|----------------------------|
| Fold | accuracy                          | f1 score | accuracy<br>unknown<br>words | accuracy<br>known<br>words | accuracy                            | f1 score | accuracy<br>unknown<br>words | accuracy<br>known<br>words |
| 0    | 90.33 %                           | 70.61 %  | 50.55 %                      | 92.14 %                    | 95.43 %                             | 66.98 %  | 61.54 %                      | 96.97 %                    |
| 1    | 87.7 %                            | 73.71 %  | 52.03 %                      | 89.81 %                    | 93.96 %                             | 65.96 %  | 70.73 %                      | 95.32 %                    |
| 2    | 89.07 %                           | 75.34 %  | 66.99 %                      | 90.26 %                    | 94.54 %                             | 66.97 %  | 66.02 %                      | 96.07 %                    |
| 3    | 87.87 %                           | 72.36 %  | 67.47 %                      | 88.75 %                    | 94.93 %                             | 66.93 %  | 63.86 %                      | 96.27 %                    |
| 4    | 87.31 %                           | 73.21 %  | 63.11 %                      | 88.54 %                    | 93.65 %                             | 66.46 %  | 72.82 %                      | 94.71 %                    |
| 5    | 91.23 %                           | 76.51 %  | 62.92 %                      | 92.47 %                    | 94.82 %                             | 68.17 %  | 69.66 %                      | 95.92 %                    |
| 6    | -                                 | -        | -                            | -                          | -                                   | -        | -                            | -                          |
| 7    | -                                 | -        | -                            | -                          | -                                   | -        | -                            | -                          |
| 8    | -                                 | -        | -                            | -                          | -                                   | -        | -                            | -                          |
| 9    | -                                 | -        | -                            | -                          | -                                   | -        | -                            | -                          |
| Avg  | -                                 | -        | -                            | -                          | -                                   | -        | -                            | -                          |

Table 4: Performances of our ingredient extractor in predicting ingredients states for fold 5 with  $\lambda = 4$