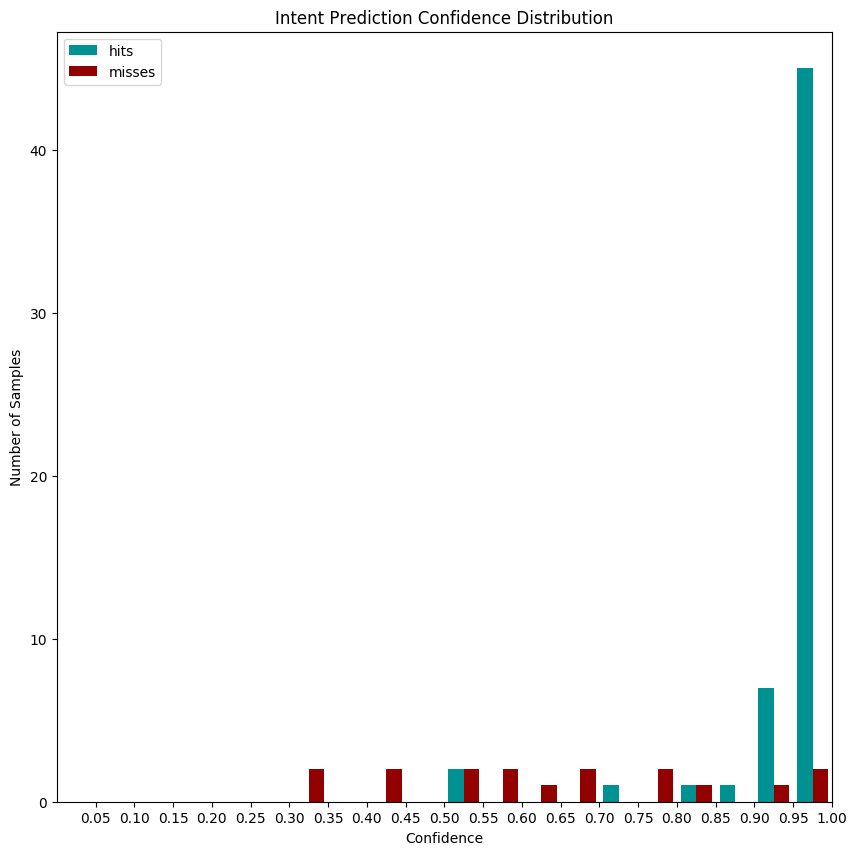
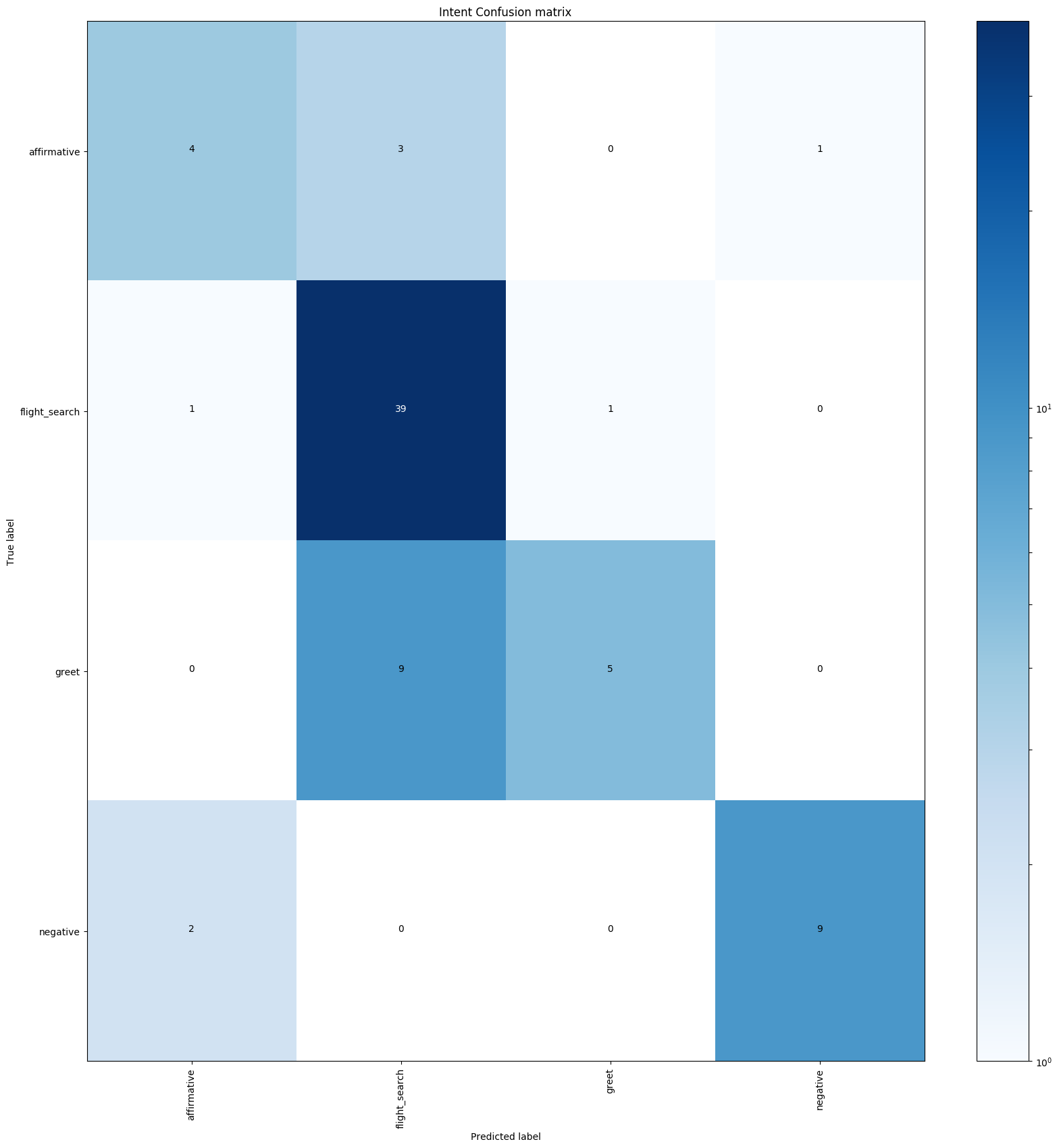
**Synthèse: Performances d’un modèle RASA NLU entraîné avec le fichier logs.json prétraité**

Après prétraitement des textes longs du fichier *logs.json* (3770 textes, cf. synthèse *Etude exploratoire)*, un modèle RASA NLU a été entraîné sur ce jeu de données, en utilisant la liste des entités spécifiées par le client.

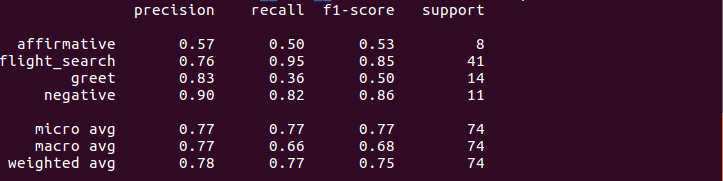
A l’issue de cet entraînement, un jeu de données de test a été mis en place pour tester et évaluer les performances du chatbot. Le jeu de données de test contient des instances de toutes les entités et tous les intents à détecter, et est constitué de structures de phrases différentes de celles vues par le chatbot durant l’entraînement.

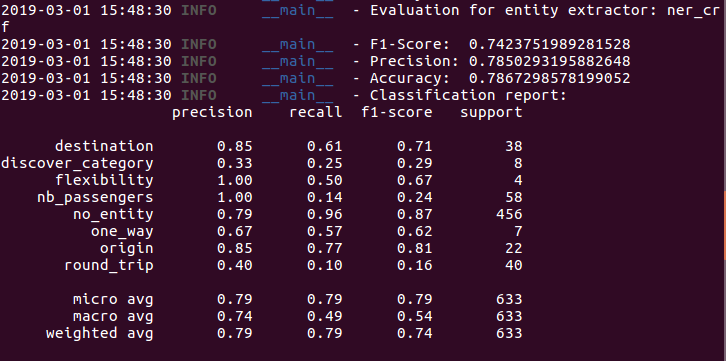
Ci-dessous le détails des performances évaluées.

**Figure 1:** Les résultats pour la classification des intents sont bons: lorsque la probabilité d’appartenance est élevée (confidence rate), les choix de classification sont (généralement) les bons.



**Figure 2:** Matrice de confusion des *intents*. Dans la globalité, les résultats de classification des *intents* sont bons, en particulier pour l’intent *flight\_search* qui nous intéresse le plus. Certains aspects restent tout de même améliorables

**Figure 3:** Précision, Recall et F1-score pour la classification des intents. Globalement, les résultats sont satisfaisants (> 70%). En particulier, la classification de l’intent flight\_search est très bonne (85 %)

**Figure 4:** Précision, recall et f1-score pour la classification des différentes entités. Le chatbot sait reconnaître et classifier correctement les entités  **destination, origin,** moyennement **flexibility, one\_way** et est faible avec les entités **nb\_passengers, discover\_category et round\_trip.** 

**Bilan:**

Ces résultats sont doublement satisfaisants. D’une part, les métriques d’évaluation sont pour la plupart des intents et des entités assez élevées: nous avons donc réussi à entraîner un chatbot qui produit des résultats similaires (voire mieux) à ceux produits par l’outil existant, en l’entrainant sur un jeu de données de faible taille, d’autre part, ces résultats nous permettent de cibler les éléments à améliorer: des exemples contenants les intents et les entités sur lesquels le chatbot se trompe doivent être présents (en grand nombre) dans le jeu de données pour le prochain entraînement. Cela est d’autant plus facile que RASA NLU permet d’obtenir cette information là facilement, sous le format ci-dessous:

[  
 {  
 "intent\_prediction": {  
 "name": "affirmative",  
 "confidence": 0.6115109941690107  
 },  
 "intent": "flight\_search",  
 "text": "vol simple tout seul "  
 },  
 {  
 "intent\_prediction": {  
 "name": "greet",  
 "confidence": 0.7702478777085815  
 },  
 "intent": "flight\_search",  
 "text": "Bonsoir ! j’vais pour zagreb"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.44588553030577965  
 },  
 "intent": "greet",  
 "text": "Salut ! est ce que ça va?"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.32097421217790967  
 },  
 "intent": "greet",  
 "text": "Bonjour monsieur"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.999970174262586  
 },  
 "intent": "greet",  
 "text": "bonsoir les amis"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.9030708770171264  
 },  
 "intent": "greet",  
 "text": "vous passez une bonne journéé monsieur?"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.4015570870025343  
 },  
 "intent": "greet",  
 "text": "slt est ce que ça va?"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.8190157713154697  
 },  
 "intent": "greet",  
 "text": "bjr"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.5105606886293947  
 },  
 "intent": "greet",  
 "text": "bjr mademoiselle"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.3407631256216708  
 },  
 "intent": "greet",  
 "text": "salut monsieur cv?"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.6859456771193869  
 },  
 "intent": "greet",  
 "text": "bonsoir mon ami ça va bien?"  
 },  
 {  
 "intent\_prediction": {  
 "name": "affirmative",  
 "confidence": 0.5475852160978402  
 },  
 "intent": "negative",  
 "text": "négative"  
 },  
 {  
 "intent\_prediction": {  
 "name": "affirmative",  
 "confidence": 0.6723480048143173  
 },  
 "intent": "negative",  
 "text": "c'est incorrect"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.7578396794695729  
 },  
 "intent": "affirmative",  
 "text": "je suis d'accord"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.9953096986753017  
 },  
 "intent": "affirmative",  
 "text": "d'accord "  
 },  
 {  
 "intent\_prediction": {  
 "name": "negative",  
 "confidence": 0.5976254062311782  
 },  
 "intent": "affirmative",  
 "text": "oui je suis d'accord"  
 },  
 {  
 "intent\_prediction": {  
 "name": "flight\_search",  
 "confidence": 0.5759712683440253  
 },  
 "intent": "affirmative",  
 "text": "d'accord c'est ça"  
 }  
]