

# Augmentation de données des agents conversationnels pour une application ressources humaines

asma.trabelsi@al-enterprise.com\*, vincent.bailleau@al-enterprise.com\*  
olivier.kreet@al-enterprise.com\*,  
andrey.leontiev@al-enterprise.com\*

\*Alcatel-Lucent Enterprise, ALE International, 32, avenue Kléber 92700 Colombes, Paris  
<https://www.al-enterprise.com/fr-fr>

**Résumé.** Les chatbots sont des agents conversationnels conçus afin d’engager des conversations textuelles avec des utilisateurs finaux. Ils ont fait l’objet de plusieurs recherches scientifiques et expérimentales dans les universités ainsi que dans l’industrie depuis l’émergence des techniques de l’intelligence artificielle et de traitement automatique du langage naturel. L’obtention des données d’apprentissage est cruciale pour la création des agents conversationnels. La majorité des chatbots nécessite la collecte des intentions ainsi que l’ensemble des expressions en langage humain, etc. Il s’agit généralement d’un processus manuel et fastidieux nécessitant des outils automatiques de génération de données d’apprentissage permettant de construire des chatbots robustes. L’une des solutions pour optimiser ce processus est l’augmentation de données. Cet article vise à étudier l’impact de l’augmentation de données sur la performance des chatbots construits sous Rasa, un outil open source pour générer les agents conversationnels. Des expérimentations ont été menées sur des données réelles. Il s’agit des données de ressources humaines permettant de construire un chatbot générique au sein de l’entreprise Alcatel-Lucent Enterprise (ALE).

## 1 Introduction

L’intelligence artificielle a pris ces dernières années une importance non négligeable dans la vie de tous les jours, que ce soit au travers des agents vocaux (Google Home, Alexa, . . . ), de l’aide à la conduite (Tesla) (Hengstler et al., 2016; Shadrin et al., 2017), ou encore des agents conversationnels (Hill et al., 2015; Luo et al., 2019). Les agents conversationnels ont connu un essor important dans plusieurs domaines : académiques (Ranoliya et al., 2017; Dibitonto et al., 2018), industriels (Okuda et Shoda, 2018; Siblini et al., 2020) et grand public afin d’automatiser nombreuses tâches dites répétitives. Le traitement de la conversation (partiellement ou intégralement) apparaît l’un des leviers libérateurs et créateurs de temps donc de richesse pour enfin recentrer la valeur du travail sur l’essentiel.

L’émergence des chatbots dans le grand public, capable d’aller plus loin dans une conversation qu’un simple échange, ainsi que l’avènement de méthode de traitement enrichie par l’apprentissage profond, permet aux entreprises d’entrevoir une application chatbot orientée business.

## L'augmentation de données pour les agents conversationnels sous RASA

Des nombreuses solutions sur le marché sont disponibles pour concevoir les chatbots telles que Dialogflow, Wit.ai, IBM Watson, Sikim, DuDy et RasaX (Bocklisch et al., 2017; Biswas, 2018; Singh et al., 2019). Alcatel-Lucent Entreprise, de son côté, développe des solutions de communication unifiée depuis de nombreuses années et propose depuis 5 ans sa plateforme Rainbow qui permet de réaliser aussi bien des communications textes, audio et/ou vidéos à destination du grand public, des entreprises et des universités. Dans ce contexte, nous avons étudié l'intégration de l'usage de chatbot dans nos produits en se basant sur l'outil de conception Rasa qui est la brique Open Source de RasaX.

Nous nous sommes focalisés sur la qualité de traitement des conversations qui est l'un des aspects les plus importants pour répondre aux contraintes des entreprises. La précision de l'interprétation du langage naturel dans ce contexte est primordiale. Ce besoin de précision est d'autant plus important qu'il peut avoir un impact sur la suite à donner à une conversation.

Bien que focalisée sur un domaine ou un secteur spécifique, la création de chatbot nécessite une phase d'apprentissage avec suffisamment de contexte et de vocabulaire, constituées à la fois des intentions et des expressions possibles, afin de s'adapter et répondre au mieux à l'utilisateur (Cheramy et François, 2019). Une même intention peut s'exprimer de différentes manières, et il est impossible pour le concepteur du chatbot de penser à toutes les options disponibles. Cette complexité va de-facto limiter et réduire la qualité d'interprétation de la machine.

La capacité d'un chatbot à être pertinent repose principalement sur la variété, la qualité et la quantité de données utilisées pour la création de l'agent intelligent (Gardent et Barahona, 2013; Wen et al., 2016). Des nombreuses solutions permettent d'atteindre cet objectif : la collecte des expressions au travers d'un panel d'utilisateur via un questionnaire récoltant une liste des questions ouvertes vis-à-vis d'une intention donnée, extraction d'un historique des mails, etc. Ces solutions nécessitent beaucoup d'intervention humaine et du temps. Une solution alternative est d'opter vers une solution d'augmentation de données. C'est un sujet qui a intéressé plusieurs chercheurs dans le domaine de traitement automatique de langage naturel afin d'enrichir d'une manière automatique la taille des données qui serviront à l'apprentissage. En effet, le choix méthodologique de recherche s'oriente vers une amélioration du corpus d'entraînement que celle de l'architecture neuronale (Xu et al., 2016; Wei et Zou, 2019; Rizos et al., 2019; Ding et al., 2020).

Dans notre papier, nous avons fait le choix d'utiliser les techniques d'augmentation sur les chatbots en utilisant le framework RASA, qui à notre connaissance une association non encore étudiée en recherche. Notre travail vise à améliorer ce framework dans un contexte d'application avec peu de données. Cette augmentation des données est basée sur des bruits appliqués pour tous les mots utilisés dans les expressions d'apprentissage ou des synonymes sémantiques ou contextuels. Le but de ce papier est d'étudier l'impact de l'augmentation de données sur la performance du chatbot, ceci en fonction du taux d'augmentation appliqué à chaque expression. Des expérimentations ont été réalisées sur une base de données questions-réponses issue des échanges quotidiens sur des questions génériques d'un service de ressources humaines, habituellement traitées manuellement et chronophage pour ces équipes.

Le reste de cet article est structuré comme suit : nous présentons dans la section 2 l'outil

Rasa (Bocklisch et al., 2017) permettant la construction des agents conversationnels et nous introduisons les approches d'augmentation de données textuelles existantes. La section 3 est dédiée à la description du processus de construction des chatbots en se basant sur Rasa tout en appliquant l'augmentation de données textuelles. Dans la section 4, nous présentons nos expérimentations afin de montrer l'impact de l'augmentation de données sur la performance des chatbots. La section 5 conclut le travail effectué.

## 2 Agents conversationnels et techniques d'augmentation de données

Dans cette section, nous soulignons l'historique des chatbots dès leurs apparitions jusqu'au aujourd'hui. Nous présentons par la suite l'outil Rasa (Bocklisch et al., 2017) qui permet de construire des agents conversationnels ainsi que les différentes techniques d'augmentation de données textuelles existantes.

### 2.1 Historique des agents conversationnels industriels

TURING (1950) a décrit ce qui est appelé le test de Turing et a introduit la notion d'agent conversationnel capable, s'il est doué d'intelligence, de tromper la personne qui échange avec lui. Même si le but initial de l'exercice était de valider si une machine est intelligente ou non, il permet d'introduire la notion d'agent conversationnel ou chatbot. A partir de 1966, des chatbots comme ELIZA (Weizenbaum, 1966) ont commencé à échanger avec leur interlocuteur avec des techniques similaires à des psychothérapies principalement en rephrasant ce que l'humain a écrit. Depuis 1990, des concours organisés chaque année cherchent à valider et se rapprocher le plus possible du test de Turing comme le Prix Loebner. C'est à cette occasion que naît ALICE, l'un des premiers chatbots à obtenir le statut d'ordinateur le plus humain en 2000, basé sur du AIML (Artificial intelligence Markup Language).

L'avènement de l'apprentissage profond popularisé grâce à des plateformes tel que Tensorflow (2015) a permis d'appliquer ses bénéfices au traitement automatique du langage naturel. Parmi les outils principaux qui permettent la réalisation de chatbot, nous pouvons citer :

- **Wit.AI** (2013), propriété de Facebook, permet de créer des chatbots simples.
- **Dialogflow**(2014), propriété de Google, permet le développement et l'intégration d'interaction texte et voix pour une intégration dans des outils tiers.
- **Amazon Lex** (2017), service AWS, nécessite un minimum d'effort de codage et de connaissance pour la création de chatbot fonctionnel.
- **IBM Watson Assistant**(2017), similaire à Dialogflow, considéré comme leader sur le marché de l'IA, orienté processus métier, permet de manière simple la création d'entité, intentions et dialogue pour la réalisation d'interfaces conversationnelles de tous types.

Il y a aussi une plateforme de développement de chatbot, qui s'appelle RASA (Bocklisch et al., 2017). Elle est gratuite et open-source et elle est similaire à Dialogflow et IBM Watson Assistant. Nous avons choisi de l'utiliser pour sa capacité d'exécution sur nos instances privées, sa flexibilité, ainsi que son ouverture.

## 2.2 L'outil RASA pour construire des agents conversationnels

Rasa, développé par Bocklisch et al. (2017), est un outil open-source permettant de construire des agents conversationnels. Il propose deux composantes principales : Rasa NLU (Natural Language Understanding) et Rasa Core. Rasa NLU a pour objectif de prédire les intentions derrière un message utilisateur et d'extraire les entités (noms propres, emails, noms de villes ou aussi des entités personnalisées). En se basant sur les techniques de l'intelligence artificielle et d'apprentissage automatique, le NLU doit être entraîné grâce à un pipeline<sup>1</sup> prenant en entrée des données d'apprentissage contenant la liste des intentions, leurs formulations possibles (les expressions), les entités ainsi que les synonymes. Une fois que le modèle est construit, il est utilisé pour déterminer la probabilité d'appartenance d'un message utilisateur à une intention et identifier les différentes entités. Le Rasa Core permet quant à lui de construire des flows de conversation en se basant sur les intentions et les entités qui sont récupérées par RASA NLU. Son rôle principal est de répondre aux messages utilisateurs par une réponse texte, un bouton ou par une action personnalisée comme par exemple l'appel à des sources de données externes ou l'appel à des APIs.

## 2.3 Techniques d'augmentation de données

L'augmentation des données est l'une des approches utilisées pour générer des données synthétiques supplémentaires en partant de données d'origine. Les méthodes d'augmentation sont souvent utilisées dans le contexte des applications de vision par ordinateur, appelé en anglais computer vision (Fawzi et al., 2016; Perez et Wang, 2017; Sallami et al., 2019). Ces techniques sont aussi performantes dans le cadre du traitement automatique de langage naturel (Kobayashi, 2018; Wei et Zou, 2019; Bari et al., 2020). Contrairement aux techniques d'augmentation pour les applications de vision par ordinateur, l'augmentation des données textuelles doit être traitée avec prudence en raison de la structure grammaticale et la complexité du texte (Coulombe, 2018). Dans ce qui suit, nous détaillons les différentes techniques d'augmentation textuelles existantes.

### 2.3.1 Différentes techniques NLP pour l'augmentation de données

Ma (2019) a mis en évidence trois types d'augmentation de données textuelles :

1. **Augmentation par caractère** : elle consiste à ajouter du bruit à un mot par l'insertion aléatoire d'un nouveau caractère, la substitution d'un caractère par un autre, l'inversion des caractères dans un mot et aussi par la suppression d'un caractère dans un mot.
2. **Augmentation par mot** : Elle a aussi pour but de rajouter du bruit dans un mot donné. Il existe plusieurs approches. Parmi ses approches, nous citons la division d'un mot en plusieurs sous-mots d'une manière aléatoire, l'inversion des mots ou la suppression d'un ou plusieurs mots dans un texte, etc. Il existe d'autres approches qui permet de remplacer un mot par un autre.

---

1. Un pipeline constitue les différents composants permettant à la fois l'entraînement du modèle NLU et le traitement de message utilisateur de manière séquentielle à fin de trouver l'intention associée et d'extraire les entités. Rasa pipeline se compose essentiellement des composants suivants : construction du dictionnaire des mots, représentation de chaque mot par un vecteur, extraction des entités et classifications des intentions

- Remplacer un mot par son synonyme tout en s'appuyant sur des bases de données open-source comme Wordnet (Miller, 1995) et PPDB (Ganitkevitch et al., 2013).
  - Remplacer un mot par un mot similaire d'une manière sémantique ou contextuelle en utilisant les techniques de plongement de mots (word embedding en anglais) (Giridhara et al., 2019). Parmi les techniques de word embedding les plus connues, nous citons TF-IDF, word2vec (Church, 2017), glove (Pennington et al., 2014), FastText (Joulin et al., 2016), Bert (Devlin et al., 2018), distilbert (Sanh et al., 2019), Roberta (Liu et al., 2019), XLNet (Yang et al., 2019), etc.
3. **Par phrase** : Remplacer une phrase par une autre en se basant soit sur les techniques de plongement de mots contextuelles comme GPT2 (Radford et al., 2019) ou XLNet (Yang et al., 2019) ou aussi de remplacer une phrase par une autre en utilisant sur les approches de résumé abstraktif (Briel, 2019).

### 2.3.2 Technique d'augmentation Chatette pour RASA

Rasa fournit un outil d'augmentation par mot permettant aussi de remplacer un mot par son synonyme. Cet outil implémente un langage DSL (appelé en anglais Domain Specific Language) afin de définir des templates et de générer un grand nombre de phrases. Il a été utilisé pour augmenter les données pour RASA NLU (Cheramy et François, 2019). En effet, il prend en entrée des blocs de texte et les entités définies à l'avance par le gestionnaire du bot pour reconstituer des phrases similaires. Cette approche nécessite donc l'intervention humaine dans la création des syntagmes et des jeux d'entités à utiliser.

## 3 Augmentation de données pour les chatbots

Nous avons conçu une plateforme permettant de développer des agents intelligents afin de transformer certaines fonctions de notre entreprise et d'améliorer notre produit Rainbow avec des briques d'intelligence artificielle. Nous nous appuyons essentiellement sur l'outil RASA (Bocklisch et al., 2017). Nous présentons dans ce qui suit l'architecture globale proposée aujourd'hui par ALE. Nous présentons aussi notre approche d'augmentation des données NLU afin d'améliorer la performance de chatbot.

### 3.1 Architecture globale adoptée

L'architecture globale de notre solution est représentée dans la Figure 1 où deux scénarios ont été identifiés. Un scénario concerne l'utilisateur final (le testeur du chatbot) et un scénario concerne le gestionnaire de l'application Web permettant la gestion et la création des bots d'une manière simple, efficace et rapide. Il s'agit d'alimenter les données concernant les deux composants à savoir le composant NLU (Intentions, Entités, Synonymes, REGEX, etc.) et le composant Core (Stories, les réponses d'intentions, les actions personnalisées). Les données sont stockées dans des serveurs de base de données. Le gestionnaire du bot peut aussi entamer la procédure d'entraînement de chatbot. Une fois que celui-ci est entraîné et déployé, un utilisateur peut commencer à communiquer avec le chatbot en passant par un connecteur. Chaque

## L'augmentation de données pour les agents conversationnels sous RASA

requête utilisateur sera émise vers le serveur Rasa où le modèle NLU va prédire l'intention derrière et le modèle Core examinera ensuite la nature de la réponse à fournir qui est une réponse de type texte ou une action personnalisée. S'il s'agit d'une action personnalisée, un serveur d'action est dédié pour traiter cette demande.

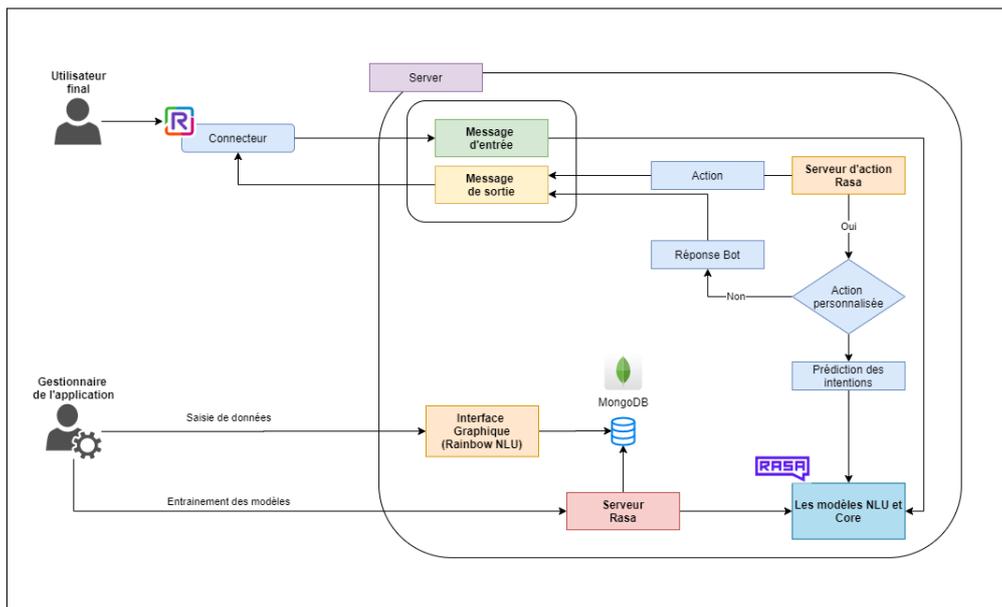


FIG. 1 – L'architecture globale proposée

### 3.2 L'augmentation de données d'apprentissage pour les chatbots

Il est important de souligner que la performance du chatbot dépend de la qualité et de la quantité de données dont nous disposons. La collecte de données pour les chatbots est une étape qui est coûteuse en terme de temps et des ressources humaines et elle présente certaines limites. En effet, il est presque impossible de couvrir toutes les expressions et les scénarios possibles. L'idée de ce papier est d'appliquer une augmentation de données servant à l'entraînement du modèle NLU afin d'améliorer la performance du chatbot. La section suivante présente notre étude expérimentale menée afin de montrer l'impact de l'augmentation de données textuelles sur la performance de chatbot et de trouver expérimentalement le meilleur compromis entre le taux d'augmentation et la performance du chatbot.

## 4 Expérimentations et résultats

Dans cette section, nous présentons notre démarche expérimentale ainsi que les résultats obtenus.

#### 4.1 Paramètres d'expérimentation

Plusieurs champs d'application ont été identifiés grâce à des besoins remontés du terrain (interne ou client). Nous avons décidé de mettre en place un bot pour les Ressources Humaines (RH). Il s'agit d'un assistant virtuel qui est capable de communiquer avec les employés d'ALE par le biais de messages instantanés. En effet, les chabots permettent à l'équipe des ressources humaines de fournir des réponses instantanées et précises pour les requêtes génériques des employés. Pour construire ce bot, une responsable RH a collecté des données réelles issues des conversations téléphoniques et des échanges par mail entre la direction RH et les salariés de notre entreprise. Les données récoltées forment un ensemble d'intentions, une réponse pour chaque intention, les expressions possibles pour chaque intention, les entités ainsi que le synonyme pour chaque entité. Le Tableau 1 décrit les détails de données RH en termes de nombre d'intentions, d'expressions, d'entités et de synonymes.

TAB. 1 – *Données RH récoltées afin de construire le chatbot*

Intentions	Expressions	Entités	Synonymes
78	652	162	479

Il est à noter que la collecte de données est une étape primordiale pour construire un chatbot mais chronophage. La collecte de données a pris environ une année pour ce contexte RH. Malgré le temps passé à collecter les données, le nombre d'expressions reste faible. Admettons que la qualité et la quantité des données impactent la performance d'un chatbot, nous avons décidé d'avoir recours à des techniques d'augmentation. Dans ce papier, deux types d'augmentation de données textuelles sont appliqués : une augmentation par Chatette et une augmentation par une technique de plongement des mots. Le premier permet de générer des ensembles de données d'entraînement pour les fichiers modèles donnés par Rasa NLU en se basant sur les entités et les synonymes. Il consiste à remplacer les mots-clés identifiés dans les expressions par leurs synonymes saisis par le gestionnaire du bot. La deuxième technique permet de trouver d'une manière automatique et sans intervention humaine les synonymes de chaque mot dans une expression mère afin de générer plusieurs expressions filles. Nous utilisons FastText car il permet de remplacer un mot par d'autres mots qui lui sont sémantiquement similaires. Afin d'appliquer l'augmentation, nous avons divisé notre base de données RH en deux parties de taille similaire : une partie pour l'apprentissage (331 expressions) et une partie pour le test (321 expressions). Nous appliquons les techniques d'augmentation sur la base d'apprentissage uniquement pour évaluer l'impact de l'augmentation sur la performance du chatbot. A notre connaissance aucun papier ne traite l'augmentation de données pour le framework RASA. Dans ce travail, nous avons privilégié de comparer différentes configurations d'approches d'augmentation en utilisant ce Framework.

#### 4.2 Résultats

Nous montrons dans un premier temps l'impact de l'augmentation sur le volume de données. La Figure 2 présente le nombre d'expressions après chaque type d'augmentation. Sur cette figure, le terme "Avec FastText- $X$ " désigne une augmentation de chaque expression

## L'augmentation de données pour les agents conversationnels sous RASA

par au plus  $X$  expressions sémantiquement équivalentes. Nous avons appliqué les techniques d'augmentation FastText-3, Chatette, une combinaison de Chatette et FastText-3, une combinaison de Chatette et FastText-5, une combinaison de Chatette et FastText-7 et une combinaison de Chatette et FastText-9. Nous avons obtenu respectivement 739, 3543, 7664, 9045, 9919 et 10887 expressions. Les résultats obtenus prouvent l'importance des techniques d'augmentation textuelles sur le volume de données. L'étape suivante consiste à étudier l'influence de cette augmentation sur la performance du chatbot. Pour ce faire, nous avons mené des expérimentations comparatives mettant en évidence la robustesse du chatbot avec les différentes techniques d'augmentation mentionnées. Les résultats empiriques sont décrits dans la Figure 3. Trois critères d'évaluation ont été utilisés, notamment le taux de classification correcte, la précision et la F-mesure. La Figure 3 présente les résultats expérimentaux de performance.

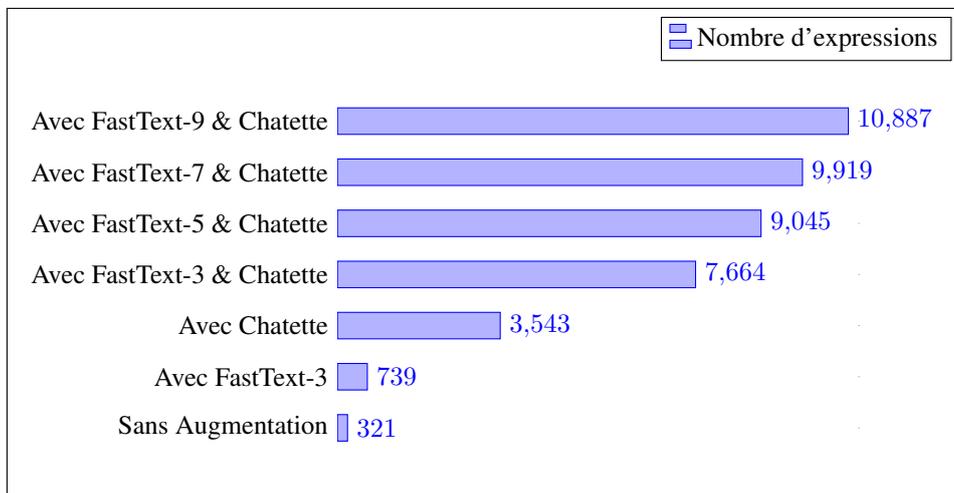


FIG. 2 – L'impact de l'augmentation de données sur le nombre des expressions

Comme indiqué dans la Figure 3, l'augmentation de données a amélioré le taux de classification correcte, la précision et la F-mesure. Plus précisément, le taux de classification correcte sans augmentation de données est égal à 51,57 % (c'est-à-dire 165 expressions parmi 321 ont été correctement classées) et il est égal à 56,6 % en appliquant la technique d'augmentation FastText-5 & Chatette (c'est-à-dire 182 expressions parmi 321 ont été correctement classées). Nous constatons la même amélioration pour la précision et la F-mesure. Nous pouvons remarquer aussi qu'une sur-augmentation de données peut réduire la performance de chatbot. Dans notre cas, nous avons obtenu un taux de classification correcte égal à 56,6%, une précision de 65,49% et une F-mesure de 55.09% pour l'augmentation avec FastText-5 & Chatette et nous avons obtenu un taux de classification correcte égal à 55,03%, une précision de 63,62% et une F-mesure de 54.49% pour l'augmentation Avec FastText-9 & Chatette. Par rapport aux résultats obtenus, il est important de souligner que dans le domaine de la reconnaissance de l'intention et de l'augmentation de données, les taux d'amélioration restent résiduels (Wei et Zou, 2019). Notre amélioration est en moyenne de 5% dans un contexte de reconnaissance de 78 intentions (l'amélioration reste prometteur). La meilleure performance est obtenu avec une combinaison de FastText-5 et de Chatette. Il est donc nécessaire de déterminer le nombre

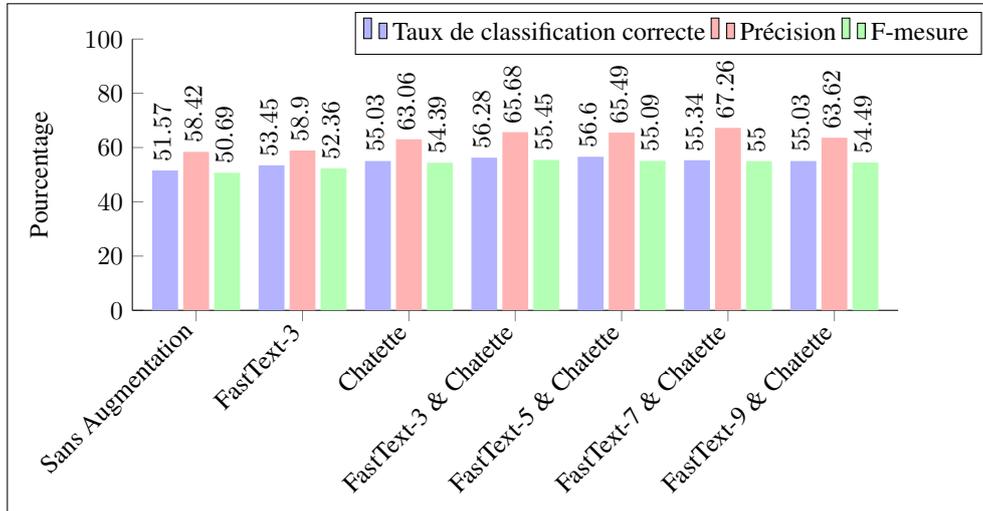


FIG. 3 – Résultats en fonction de taux de classification correcte, Précision et F-mesure

optimal de données à augmenter pour avoir les meilleures performances.

Pour le calcul, nous avons à notre disposition une machine dotée d'un processeur Intel(R) Xeon(R) E-2274G CPU @ 4.00GHz avec 64 Go de RAM afin d'effectuer l'augmentation de données et nous avons utilisé une machine avec un processeur Intel(R) Xeon(R) CPU E3-1270 v6 @ 3.80GHz sous kubernetes pour l'entraînement des modèles. Le temps de calcul nécessaire pour l'augmentation de données et pour l'apprentissage est donné dans le tableau ci-dessous.

TAB. 2 – Temps de calcul

Types d'augmentation	Avec FastText-3	Avec Chatette	Avec FastText-3 & Chatette	Avec FastText-5 & Chatette	Avec FastText-7 & Chatette	Avec FastText-9 & Chatette
Temps d'augmentation de données (min)	14	2	15	25	37	48
Temps d'apprentissage de données (min)	8	14	32	37	45	69
Total (min)	22	16	47	62	82	117

Comme indiqué dans le Tableau 2, le temps de calcul augmente avec la quantité de données. Un modèle construit avec Fast-Text-7 et Chatette nécessite 1h et 22 min du calcul sur une machine dotée d'un CPU. Il est donc important d'opter vers des machines avec des GPUs afin de réduire le temps de calcul.

## 5 Conclusion

Le but de ce papier est d'étudier l'impact de l'augmentation de données sur la performance des chatbots. Nous avons effectué des études expérimentales en utilisant une base de données question-réponse générique issue de la RH ALE. Les résultats expérimentaux ont prouvé une amélioration de la robustesse du chatbot suite à l'augmentation de données. Ils ont prouvé aussi qu'il est crucial de trouver le taux d'augmentation optimal. Aujourd'hui, l'ajustement du paramètre  $k$  pour FasText a été effectué d'une manière manuelle. Dans nos futurs travaux, nous essaierons d'introduire des nouvelles approches afin d'automatiser la recherche du paramètre optimal pour l'augmentation de données textuelles.

## Références

- Bari, M. S., M. T. Mohiuddin, et S. Joty (2020). Multimix : A robust data augmentation strategy for cross-lingual nlp. *arXiv preprint arXiv :2004.13240*.
- Biswas, M. (2018). Wit. ai and dialogflow. In *Beginning AI Bot Frameworks*, pp. 67–100. Springer.
- Bocklisch, T., J. Faulkner, N. Pawlowski, et A. Nichol (2017). Rasa : Open source language understanding and dialogue management. *arXiv preprint arXiv :1712.05181*.
- Briel, A. (2019). Abstractive summarization for data augmentation, <https://towardsdatascience.com/abstractive-summarization-for-data-augmentation-1423d8ec079e>.
- Cheramy, J. et T. François (2019). La génération de paraphrases de données d'entraînement pour une meilleure classification des intents d'un chatbot créé avec rasa, <https://dial.uclouvain.be/memoire/ucl/fr/object/thesis>
- Church, K. W. (2017). Word2vec. *Natural Language Engineering* 23(1), 155–162.
- Coulombe, C. (2018). Text data augmentation made simple by leveraging nlp cloud apis. *arXiv preprint arXiv :1812.04718*.
- Devlin, J., M.-W. Chang, K. Lee, et K. Toutanova (2018). Bert : Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv :1810.04805*.
- Dibitonto, M., K. Leszczynska, F. Tazzi, et C. M. Medaglia (2018). Chatbot in a campus environment : design of lisa, a virtual assistant to help students in their university life. In *International Conference on Human-Computer Interaction*, pp. 103–116. Springer.
- Ding, B., L. Liu, L. Bing, C. Kruengkrai, T. H. Nguyen, S. Joty, L. Si, et C. Miao (2020). An effective data augmentation method for low-resource tagging tasks. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 6045–6057.
- Fawzi, A., H. Samulowitz, D. Turaga, et P. Frossard (2016). Adaptive data augmentation for image classification. In *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3688–3692. Ieee.
- Ganitkevitch, J., B. Van Durme, et C. Callison-Burch (2013). Ppdb : The paraphrase database. In *Proceedings of the 2013 Conference of the North American Chapter of the Association*

- for Computational Linguistics : Human Language Technologies*, pp. 758–764.
- Gardent, C. et L. M. R. Barahona (2013). Using paraphrases and lexical semantics to improve the accuracy and the robustness of supervised models in situated dialogue systems. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 808–813.
- Giridhara, P. K. B., C. Mishra, R. K. M. Venkataramana, S. S. Bukhari, et A. Dengel (2019). A study of various text augmentation techniques for relation classification in free text. *IC-PRAM* 3, 5.
- Hengstler, M., E. Enkel, et S. Duelli (2016). Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change* 105, 105–120.
- Hill, J., W. R. Ford, et I. G. Farreras (2015). Real conversations with artificial intelligence : A comparison between human–human online conversations and human–chatbot conversations. *Computers in human behavior* 49, 245–250.
- Joulin, A., E. Grave, P. Bojanowski, M. Douze, H. Jégou, et T. Mikolov (2016). Fasttext. zip : Compressing text classification models. *arXiv preprint arXiv :1612.03651*.
- Kobayashi, S. (2018). Contextual augmentation : Data augmentation by words with paradigmatic relations. *arXiv preprint arXiv :1805.06201*.
- Liu, Y., M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, et V. Stoyanov (2019). Roberta : A robustly optimized bert pretraining approach. *arXiv preprint arXiv :1907.11692*.
- Luo, X., S. Tong, Z. Fang, et Z. Qu (2019). Frontiers : Machines vs. humans : The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science* 38(6), 937–947.
- Ma, E. (2019). Nlpaug, <https://github.com/makcedward/nlpaug>.
- Miller, G. A. (1995). Wordnet : a lexical database for english. *Communications of the ACM* 38(11), 39–41.
- Okuda, T. et S. Shoda (2018). Ai-based chatbot service for financial industry. *Fujitsu Scientific and Technical Journal* 54(2), 4–8.
- Pennington, J., R. Socher, et C. D. Manning (2014). Glove : Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543.
- Perez, L. et J. Wang (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv :1712.04621*.
- Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, et I. Sutskever (2019). Language models are unsupervised multitask learners. *OpenAI Blog* 1(8), 9.
- Ranoliya, B. R., N. Raghuwanshi, et S. Singh (2017). Chatbot for university related faqs. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1525–1530. IEEE.
- Rizos, G., K. Hemker, et B. Schuller (2019). Augment to prevent : short-text data augmentation in deep learning for hate-speech classification. In *Proceedings of the 28th ACM International*

- Conference on Information and Knowledge Management*, pp. 991–1000.
- Sallami, M. M., M. I. Khedher, A. Trabelsi, S. Kerboua-Benlarbi, et D. Bettebghor (2019). Safety and robustness of deep neural networks object recognition under generic attacks. In *International Conference on Neural Information Processing*, pp. 274–286. Springer.
- Sanh, V., L. Debut, J. Chaumond, et T. Wolf (2019). Distilbert, a distilled version of bert : smaller, faster, cheaper and lighter. *arXiv preprint arXiv :1910.01108*.
- Shadrin, S. S., O. O. Varlamov, et A. M. Ivanov (2017). Experimental autonomous road vehicle with logical artificial intelligence. *Journal of advanced transportation 2017*.
- Siblini, W., C. Pasqual, A. Lavielle, et C. Cauchois (2020). Système de question-réponse multilingue appliqué aux agents conversationnels. In *EGC*, pp. 333–340.
- Singh, A., K. Ramasubramanian, et S. Shivam (2019). Introduction to microsoft bot, rasa, and google dialogflow. In *Building an Enterprise Chatbot*, pp. 281–302. Springer.
- TURING, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind LIX*(236), 433–460.
- Wei, J. et K. Zou (2019). Eda : Easy data augmentation techniques for boosting performance on text classification tasks. *arXiv preprint arXiv :1901.11196*.
- Weizenbaum, J. (1966). Eliza—a computer program for the study of natural language communication between man and machine. *Commun. ACM* 9(1), 36–45.
- Wen, T.-H., D. Vandyke, N. Mrksic, M. Gasic, L. M. Rojas-Barahona, P.-H. Su, S. Ultes, et S. Young (2016). A network-based end-to-end trainable task-oriented dialogue system. *arXiv preprint arXiv :1604.04562*.
- Xu, Y., R. Jia, L. Mou, G. Li, Y. Chen, Y. Lu, et Z. Jin (2016). Improved relation classification by deep recurrent neural networks with data augmentation. *arXiv preprint arXiv :1601.03651*.
- Yang, Z., Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, et Q. V. Le (2019). Xlnet : Generalized autoregressive pretraining for language understanding. In *Advances in neural information processing systems*, pp. 5753–5763.

## Summary

Chatbots are conversational agents designed to engage in text-based conversations with end-users. They have been the subject of several scientific and experimental researches in universities as well as in the industry since the emergence of artificial intelligence techniques and natural language processing. Obtaining the right example of training data is crucial for the creation of chatbots. Most chatbots require the collection of intentions as well as expressions in human language. This is generally a manual and tedious process and hence the need for automatic learning data volume generation tools to build robust chatbots. One of the solutions that exists is the increase of data in natural language. This article aims to study the impact of increased data on chatbot performance. To do this, we carried out experiments on real data. This is the human resources data used to build a generic chatbot within our company Alcatel-Lucent Enterprise (ALE).