

Reconnaissance de la parole : Microsoft atteint la parité homme-machine

● En matière de reconnaissance de la parole, Microsoft a fait d'énormes avancées. Pour la première fois, un ordinateur peut reconnaître les mots d'un flux de parole conversationnelle de la même façon que le fait un interlocuteur humain.

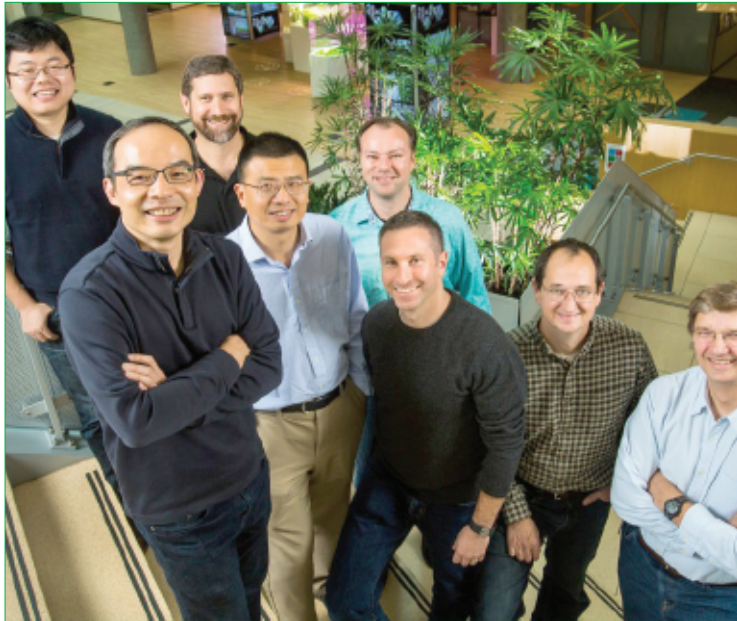
Par Farid Farah

Les dix dernières années témoignent des énormes progrès réalisés dans l'amélioration de la fiabilité des systèmes de traitement automatique de la parole. C'est Microsoft qui a cette fois réalisé une percée exceptionnelle dans ce domaine dont la plupart des travaux sont au stade de la recherche. Une de ses équipes de chercheurs a créé une technologie vocale visant à permettre à une machine de reconnaître les mots de la parole conversationnelle de la même manière qu'un interlocuteur humain. En effet, des chercheurs de Microsoft viennent de mettre en évidence une nouvelle technique de reconnaissance de la parole conversationnelle qui permet d'atteindre les mêmes performances d'un interlocuteur humain. Inspirée du concept des réseaux de neurones, cette méthode consiste à répertorier les mots synonymes phonologiquement différents dans la même classe des vecteurs analytiques. Dans un article paru le 17 octobre dernier, et téléchargeable sur le site arxiv.org (archives ouvertes de publications scientifiques) de la Cornell University Library, une équipe de chercheurs de Microsoft relevant du laboratoire de la recherche en intelligence

artificielle a réalisé un système de reconnaissance automatique de la parole en conversation qui s'approche des performances de reconnaissance d'un locuteur humain. Durant la phase des tests définitifs, les ingénieurs de la firme américaine ont mesuré un taux d'erreur de mots WER (Word Error Rate) de 5.9 % contre 6.3 % il y a encore un mois. Selon les spécialistes de la reconnaissance vocale, un taux de 5.9 % est approximativement identique à celui des interlocuteurs humains qui ont été invités à transcrire le même contenu conversationnel. Pour le géant du logiciel PC, il s'agit du taux d'erreur le plus bas jamais enregistré dans le secteur de l'industrie de la reconnaissance de la parole standard. «Nous avons atteint la parité humaine», a déclaré Xuedong Huang, responsable des sciences de la parole chez Microsoft.

Des décennies d'efforts

Les résultats des travaux de recherche des ingénieurs de Microsoft ont largement dépassé les attentes de leurs encadrement : pour la première fois de l'histoire de l'informatique, un ordinateur ou un autre équipement terminal peut reconnaître les mots composant un flux de parole conversationnelle de la même façon qu'un interlocuteur humain le fait. Cette innovation historique s'est finalement réalisée après des décennies d'efforts dans le développement des systèmes de reconnaissance de la parole performants. Les travaux de recherche ont débuté durant les années 1970 quand l'agence DARPA (Defense Advanced Projects Agency), un organisme américain chargé de la promotion



Photos : DR

des technologies multidisciplinaires au profit de la sécurité nationale du pays, a été créée. Au fil des années, plusieurs compagnies technologiques et organismes de recherche se sont joints aux projets de la DARPA. Cette réalisation technologique aura de larges impacts sur les produits IT dont le système logique sera considérablement amélioré par la reconnaissance de la parole. Les appareils de divertissement grand public comme la Xbox, les outils d'accessibilité tels que la transcription instantanée de la parole en texte et les assistants numériques personnels comme Cortana seront plus adaptés aux utilisateurs. «Pour nous, une telle performance, apportera plus d'intelligence à notre assistant numérique personnel Cortana», a affirmé Harry Shum, vice-président exécutif de Microsoft qui dirige le pôle de

recherche dans le domaine de l'intelligence artificielle. Par ailleurs, cette innovation ne mènera pas l'ordinateur vers une reconnaissance parfaite de chaque mot articulé. Il en est de même pour l'être humain.

Plus d'intelligence pour Cortana

L'exploit des chercheurs de Microsoft réside dans le fait que l'outil de reconnaissance de la parole maîtrise parfaitement la vitesse d'articulation des mots du contenu de la parole conversationnelle. A l'image d'un interlocuteur humain, la machine peut par exemple entendre un «ont» à la place d'un «est» si la vitesse d'articulation est élevée. Les chercheurs affirment dans leur article

que le secret du triomphe technologique de leur système est l'utilisation systématique des réseaux de neurones convolutifs et récurrents LSTM (Long Short Term Memory) dans toutes les parties du système. Cette nouvelle méthode de modélisation des signaux de la parole se base sur le principe de représenter les mots de la parole par des vecteurs continus dans l'espace. Elle permet de répertorier les mots synonymes et phonologiquement différents dans le même groupe vectoriel. Par exemple, les mots anglais «rapid» et «speedy» sont classés dans un même groupe de vecteurs. Le processus de la généralisation des modèles de reconnaissance de la parole par mots sera alors facilité. Le modèle acoustique LSTM permet aux réseaux de neurones d'utiliser de grandes quantités de données appelées groupes d'entraînement pour apprendre aux systèmes informatiques la meilleure manière de reconnaître les formes à partir des données d'entrées telles que les sons et images. Et pour lever le défi de la parité humaine-machine, les chercheurs ont également utilisé la boîte à outils cognitifs de Microsoft CNTK (Computational Network Toolkit). Disponible sur GitHub via une licence open source, ce Toolkit a la capacité de traiter rapidement les algorithmes d'intensification de l'apprentissage en utilisant plusieurs ordinateurs tournant sous une même puce appelée unité de traitement graphique.

F. F.

TAGS Microsoft, Reconnaissance, Parole, DARPA, GitHub, Computational Network Toolkit, Cornell University Library, Cortana, Intelligence artificielle.

APRÈS SON SUCCÈS SUR FACEBOOK

«Info Trafic Algérie» se dote d'un site web

● Info Trafic Algérie (ITA, pour les intimes), une des pages algériennes les plus suivies sur Facebook, avec près de 100 000 membres, s'est dotée depuis quelques jours d'un site web. Un «passage obligé» et «revendiqué» par les utilisateurs, explique Meriem Khelil, l'initiatrice de ce projet.

Abdelkader Zahar

«La RN12 sera fermée par les habitants du village Bethlou demain 23/10. Prenez vos précautions !» Ceci n'est pas une information parue dans la presse ou un SMS du ministère des Travaux publics, mais une alerte lancée sur le site web «Info Trafic Algérie» publiée la veille à 18h24 dans la rubrique «Rana Rayhine». L'info parue initialement sur la page Facebook de ITA, grâce au réseau de conducteurs qui y est abonné, a été relayée immédiatement vers le site.

Après trois années d'existence sur le réseau social préféré des Algériens, Info Trafic Algérie, la page FB consacrée à l'actualité de la circulation routière en Algérie, passe en mode site web. Une «étape nécessaire, voulue et revendiquée par de très nombreux membres de la page ITA sur Facebook», affirme Meriem Khelil. Le nouveau site a été lancé le 22 octobre 2016 qui coïncide avec la «Journée nationale de la presse algérienne». L'idée de «compléter»

ITA par un site internet «remonte à plusieurs mois», affirme M^{me} Khelil. Et celle de lancer une «application mobile» est dans l'agenda. Lancée en mai 2013, la page Facebook ITA compte 99 828 abonnés, alors que le groupe fermé éponyme en est à plus de 109 000 membres. Sur ces deux «adresses» de Facebook, les intervenants sont très nombreux. Et sur les dizaines de publications quotidiennes, il arrive que certaines, parmi elles, ne soient pas avérées, ou alors non d'actualité (des images évoquant des événements dépassés). Sur le site web ITA, les informations sont contrôlées. «Toutes les publications sont vérifiées et/ou complétées par les modérateurs», affirme Meriem Khelil.

Marque déposée

A la différence de la page et du groupe Facebook, la plateforme web Info Trafic Algérie comporte plusieurs rubriques. La plus importante est «Rana Rayhine» (nous allons partir), le fil d'actualité où sont publiées les alertes de circulation routière (engorgements, fluidité, routes barrées, accidents) avec la mention du lieu et de l'heure pour plus de précision. Toujours en page d'accueil, le site reprend des articles de presse relatifs aux transports, circulation routière, permis de conduire et aux accidents. Cadre dans le secteur des télécoms, Meriem Khelil parcourt, «chaque matin avant 6h30», les sites de la presse écrite pour «collecter des infos» en relation avec l'objet d'Info Trafic Algérie. Le reste de l'actualité routière est



puisé dans les publications des membres de la page Facebook. Le site propose six autres rubriques : «RP 360°» (revue de presse quasi quotidienne), «Lois et réglementation» (en relation avec le secteur des transports et circulation routière), «Interview» (encore vide), «Témoignages» de conducteurs sur l'état des routes et autres soucis rencontrés en conduite, «Astuces automobiles» (nettoyage, pneumatique), et des «Quizz Auto».

Pour le moment, «Info Trafic Algérie» fonctionne en «site fédérateur». «Je n'ai pas de journalistes pour rédiger les informations ou les vérifier. Je me fais aider par mon mari», explique Meriem Khelil. Conçu et hébergé par kdconcept, le site «est en constante amélioration en termes de contenu

et d'affichage», affirme-t-elle. A noter que le site est «entièrement financé» par son initiatrice. «Les annonceurs sont les bienvenus dans les espaces publicitaires réservés à cet effet», dit-elle. Avant de préciser que le nom «Info Trafic Algérie» est une «marque déposée», insiste M^{me} Khelil. Une précision qui s'impose quand on sait qu'un site «concurrent», portant le même nom, créé le 16 septembre 2016, «puise toutes ses informations de la page Facebook de ITA», ajoute Meriem Khelil.

A. Z.

TAGS Info Trafic Algérie, Facebook, Meriem Khelil, Application mobile, Site web, Circulation routière