

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 8 ماي 1945
École Supérieure en Informatique
8 Mai 1945 Sidi Bel Abbès



MÉMOIRE

En vue de l'obtention du diplôme de **Master**
Filière: **Informatique**
Spécialité: **Ingénierie des Systèmes Informatiques (ISI)**

Thème

SEMANTIC TEXTUAL SIMILARITY IN CHAT BOTS : COMPARATIVE
STUDY

Présenté par:

Hammouda Selouane

Soutenu le : **03/07/2023** devant le jury composé de :

dr. KAZI TANI Mohammed Yassin	President
dr. CHAIB Souleyman	Encadrant
dr. BEKKOUCHE Mohammed	Examineur

Année Universitaire : 2022-2023

ABSTRACT

Chatbots are valuable tools in numerous applications across industries, they offer several advantages to boost the overall functionality and performance.

In literature, there are several ways to implement a chat bot, however this research paper focuses on incorporating semantic textual similarity model in chatbots. Furthermore, the thesis presents a comparative study in modeling semantic similarity aiming to identify the state-of-the-art models in the field.

Key words: Chatbots, Semantic Textual Similarity, Word Embedding, Natural Language Processing, Deep Learning, Long Short-Term Memory, Convolutional Neural Network, Recurrent Neural Network, Transformers.

Les chatbots sont des outils précieux dans de nombreuses applications à travers les industries, ils offrent plusieurs avantages pour augmenter la fonctionnalité et des performances globales.

Dans la littérature, il existe plusieurs façons de mettre en œuvre un robot de clavardage, mais ce document de recherche se concentre sur l'intégration d'un modèle de similitude textuelle sémantique dans les robots de clavardage. En outre, la thèse présente une étude comparative sur la modélisation de la similarité sémantique visant à identifier les modèles de état de l'art dans le domaine.

Mots clés: Chatbots, Similitude Sémantique Textuelle, Intégration de Mots, Traitement Naturel du Langage, Apprentissage Profond, Mémoire à Court Terme, Réseau de Neurones Convolutionnels, Réseau de Neurones Récurrents, Transformateurs.

مُلخَص

تعد الدردشة الآلية أداة قيمة في تطبيقات البرامج ، فهي توفر العديد من المزايا التي تحسن من وظائف وأداء النظم بشكل عام.

في البحوث العلمية، ذكرت عدة طرق لتنفيذ روبوت الدردشة، الا ان هذه المذكرة ستركز بشكل اساسي على دمج نموذج التشابه النصي الدلالي في روبوتات الدردشة. بالاضافة الى ذلك، ستقدم الاطروحة دراسة مقارنة في نمذجة التشابه الدلالي بهدف تحديد أحدث النماذج في هذا المجال.

كلمات مفتاحية: روبوتات الدردشة، التشابه النصي الدلالي، تضمين الكلمات، معالجة اللغة الطبيعية، التعلم العميق، الذاكرة قصيرة المدى، الشبكة العصبية التلافيفية، الشبكة العصبية المتكررة، المحولات.

THANKS

Endless thanks to **Allah Almighty** for constantly blessing me with much needed strength and persistence to carry on and do my best in favor of this project to be completed despite all the bumps on the road.

I would like to express my deepest gratitude for my supervisors **Chaib Souleyman** for his unwavering support and insightful feedback throughout the entire process of this research.

It is worth casting light on the fact that the evaluation and comments of each and every member of the jury are much appreciated in advance.

To **BGB Team, Fethi Chennaoui, Zeyneb Ghezal** and **Firaz Benmokadem** thank you for your hospitality. Your warm welcome and graciousness made the time I spent as an intern truly memorable.

My greatest appreciation is due to BGB IT Team: **Madji Youcef, Mokrane Aoulaiche** and **Amrine Moussab Amine**. I am profoundly grateful for the opportunity to learn from you. Your expertise and willingness to share your knowledge have contributed heavily to my personal and professional growth.

To my lovely **parents**, I cannot thank you enough for never once quitting on me. The ceaseless encouragement as well as the amount of faith both of you have put in me were my source of motivation to reach new heights.

Special thanks go to my supportive friends: **Manal, Hafsa, Warda, Fella, Ikram, Dounia, Kouather, Malika, Ramzi, Mahmoud, Khaled** and all **GDG baraiim**. I am very appreciative of the sense of safety you have provided me with as well as keeping me grounded during stressful times.

DEDICATION

I sincerely dedicate this work to :

My beloved parents, who embraced me with their love and care in the darkest times. Your everlasting support, encouragement, and sacrifices have been the foundation of my journey. I am forever in your debt.

*My BGB supervisor **Youcef Madji** Your mentorship has not only shaped my research but inspired me to envision new perspectives and further comprehend the subject matter, as well.*

*My loveliest bestfriend **Manal Bouriba** Your endless support, cheer, and belief in my abilities have meant the world to me. Your presence during the highs and lows of this academic endeavor has been a source of strength and motivation. Thank you for always being there and for reminding me of the importance of balance, laughter, and friendship.*

NLP Natural Language Processing.

POS Part Of Speech .

CNN Convolutional Neural Network.

RNN Recurrent Neural Network .

LSTM Long Short-Term Memory.

STS Semantic Textual Similarity.

NER Named Entity Recognition.

BERT Bidirectional Encoder Representations from Transformers.

SBERT Sentence BERT.