

# Washington Cunha, Ph.D. Student

✉ washingtoncunha@dcc.ufmg.br

in <https://www.linkedin.com/in/washington-l-m-cunha/>

📞 +5531992179467

🐙 <https://www.github.com/waashk>

## Research Interest

Washington Cunha    I'm a Ph.D. Student in the Machine Learning and Databases Laboratory (LBD), advised by teacher Marcos Goncalves at Federal University of Minas Gerais. I'm working on Deep Learning Text Classification (ATC) and Sentiment Analysis (SA) models. More specifically, my main research goal is to study instance selection approaches in order to reduce the cost (time consumption) of deep learning ATC and SA models while keeping the effectiveness.

## Education

- 2019 – Actual    **Ph.D. Student**, Federal University of Minas Gerais.  
Project Title: *On the Role of Instance Selection Methods for Text Classification*
- 2018 – 2019    **Master Degree in Computer Science** in Federal University of Minas Gerais  
Thesis title: *Extended pre-processing pipeline for text classification: On the role of meta-feature representations, sparsification and selective sampling.*
- 2014 – 2017    **Bachelor Degree in Computer Science** in Federal University of São João del-Rei.  
Thesis title: *A Feature-oriented Sentiment Rating for Mobile App Reviews.*

## Research Publications

### Journal Articles

- 1 On the cost-effectiveness of neural and non-neural approaches and representations for text classification. (2021). *IPeM – Imp. Fac.*: 6.222. doi:10.1016/j.ipm.2020.102481
- 2 Stroke outcome measurements from electronic medical records: Cross-sectional study on the effectiveness of neural and nonneural classifiers. (2021). *JMIR Med Inform.* doi:10.2196/29120
- 3 Extended pre-processing pipeline for text classification: On the role of meta-feature representations, sparsification and selective sampling. (2020). *IPeM.* doi:10.1016/j.ipm.2020.102263

### Conference Proceedings

- 1 CluHTM - Semantic Hierarchical Topic Modeling based on CluWords. (2020), In *ACL'20 – h5-index: 135.* doi:10.18653/v1/2020.acl-main.724
- 2 “Keep It Simple, Lazy” – MetaLazy: A New MetaStrategy for Lazy Text Classification. (2020), In *CIKM'20 – h5-index: 59.* doi:10.1145/3340531.3412180
- 3 CluWords: Exploiting Semantic Word Clustering Representation for Enhanced Topic Modeling. (2019), In *WSDM'19 – h5-index: 61.* doi:10.1145/3289600.3291032
- 4 A Feature-oriented Sentiment Rating for Mobile App Reviews. (2018), In *The world wide web conference WWW'18 – h5-index: 90.* doi:10.1145/3178876.3186168
- 5 Semantically-Enhanced Topic Modeling. (2018), In *CIKM'18 – h5-index: 59.* doi:10.1145/3269206.3271797

## Skills

|           |  |
|-----------|--|
| Coding    | Python, C, C++, CUDA, R, AWK, Shellscript, $\LaTeX$ , PowerBi, Tableau   |
| Research  | Machine Learning and Data Mining, focusing in Automatic Text Classification, Feature Engineering, Sentiment Analysis and Topic Modeling. |
| Libraries | <b>Python:</b> Scikit-learn, Numpy, Pandas, Pytorch, Keras, Scipy, Mxnet, Matplotlib, Seaborn.   |
| Dev       | AWS, Linux, Git, Docker, Jupyter-notebook, Google Colab.   |
| Languages | Portuguese and English.  |
| Misc.     | Academic research, teaching, training, consultation, $\LaTeX$ typesetting and publishing.  |

## Miscellaneous Experience

### Awards and Achievements

- 2021 **Honorable Mention** in the Masters Theses Contest of the Brazilian Database Symposium. **CTDBD – SBBD’21**
- 2019 **Among of 10 best Brazilian Scientific Initiations** Research selected by Brazilian Computer Society (As co-advisor). **CTIC’19 – SBC.**

### Advanced Courses

- 2020 Natural Language Processing - Deep Learning Algorithms
- 2019 Deep Learning Algorithms
- Information Theory
- 2018 Information Retrieval and Social Computing
- Quantitative Methods of Experimental Research in Computer Science
- Machine Learning
- Fundamentals of Statistics for Data Science

### Certification

- 2020 Coursera Specialization: Deep Learning - Coursera
- Coursera Specialization: Applied Data Science with Python - Coursera
- Coursera Specialization: AWS Fundamentals - Coursera
- Practical Project Management - Udemy
- 2019 Fundamentals of Accelerated Computing with CUDA C/C++ - NVIDIA

## Academic Service

- 2021 Association for Computational Linguistics (**ACL**) - Reviewer
- The Conference on Empirical Methods in Natural Language Processing (**EMNLP**) - Reviewer
- The Conference on Information and Knowledge Management (**CIKM**) - Sub-Reviewer
- The European Conference on Information Retrieval (**ECIR**) - Sub-Reviewer
- Association for Computational Linguistics (**ACL**) Mentorship - Participant
- WebMedia Conference - Volunteer - Squad Leader
- ACM and SBC Membership