Secondary Bibliography: Papers Surveyed

Surveyed Papers

Note: The list of surveyed papers starts with **ID = 2**, corresponding to the numbering in the associated Excel file.
The Excel file can be accessed at the following link: `https://docs.google.com/spreadsheets/d/10HcCSnKe-HFNg86B6cxzB3pL4GjZO6pcS_RQbNh94Yc/edit?usp=sharing`

Paper ID: 2
Authors: Jen-tse Huang, Jianping Zhang, Wenxuan Wang, Pinjia He, Yuxin Su, Michael R. Lyu
Title: AEON: A Method for Automatic Evaluation of NLP Test Cases
Venue: Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis
Pages: 202–214
Year: 2022
DOI: `https://doi.org/10.1145/3533767.3534394`

Paper ID: 3
Authors: Fangqing Liu, Han Huang, Zhongming Yang, Zhifeng Hao, Jiangping Wang
Title: Search-based algorithm with scatter search strategy for automated test case generation of NLP toolkit
Venue: IEEE Transactions on Emerging Topics in Computational Intelligence
Details: 5(3):491–503
Year: 2019
DOI: `https://doi.org/10.1109/TETCI.2019.2914280`

Paper ID: 4
Authors: Rotem Dror, Gili Baumer, Segev Shlomov, Roi Reichart
Title: The hitchhiker's guide to testing statistical significance in natural

language processing
Venue: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)
Details: Pages 1383–1392
Year: 2018
DOI: https://doi.org/10.18653/v1/P18-1128

Paper ID: 5
Authors: Brihi Joshi, Aaron Chan, Ziyi Liu, Shaoliang Nie, Maziar Sanjabi, Hamed Firooz, Xiang Ren
Title: Er-test: Evaluating explanation regularization methods for NLP models
Venue: arXiv preprint
Details: arXiv:2205.12542
Year: 2022
Link: https://arxiv.org/pdf/2205.12542

Paper ID: 6
Authors: Marco Tulio Ribeiro, Scott Lundberg
Title: Adaptive testing and debugging of NLP models
Venue: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)
Details: Pages 3253–3267
Year: 2022
DOI: https://doi.org/10.18653/v1/2022.acl-long.230

Paper ID: 7
Authors: Ruibo Wang, Jihong Li
Title: Bayes test of precision, recall, and F1 measure for comparison of two natural language processing models
Venue: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics
Details: Pages 4135–4145
Year: 2019
Link: https://aclanthology.org/P19-1405.pdf

Paper ID: 8
Authors: Guanqun Yang, Mirazul Haque, Qiaochu Song, Wei Yang, Xueqing Liu
Title: Testaug: A framework for augmenting capability-based NLP tests
Venue: arXiv preprint
Details: arXiv:2210.08097

Paper ID: 9
Authors: Yaoqing Yang, Ryan Theisen, Liam Hodgkinson, Joseph E. Gonzalez, Kannan Ramchandran, Charles H. Martin, Michael W. Mahoney
Title: Test accuracy vs. generalization gap: Model selection in NLP without accessing training or testing data
Venue: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining
Details: Pages 3011–3021
Year: 2023
DOI: https://doi.org/10.1145/3580305.3599518

Paper ID: 10
Authors: Samson Tan, et al.
Title: Reliability Testing for Natural Language Processing Systems
Venue: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)
Details: Pages 3240–3252
Year: 2021
DOI: https://doi.org/10.48550/arXiv.2105.02590

Paper ID: 11
Authors: Yaoqing Yang, Ryan Theisen, Liam Hodgkinson, Joseph E. Gonzalez, Kannan Ramchandran, Charles H. Martin, Michael W. Mahoney
Title: Evaluating natural language processing models with generalization metrics that do not need access to any training or testing data
Venue: arXiv preprint
Details: arXiv:2202.02842
Year: 2022
DOI: https://doi.org/10.48550/arXiv.2202.02842

Paper ID: 12
Authors: Konstantinos Kogkalidis, Stergios Chatzikyriakidis, Eirini Chrysovalantou Giannikouri, et al.
Title: OYXOY: A Modern NLP Test Suite for Modern Greek
Venue: arXiv preprint
Details: arXiv:2309.07009
Year: 2023
DOI: https://doi.org/10.48550/arXiv.2309.07009

Paper ID: 13
Authors: Carlos Aspillaga, Andrés Carvallo, Vladimir Araujo
Title: Stress test evaluation of transformer-based models in natural language understanding tasks
Venue: arXiv preprint
Details: arXiv:2002.06261
Year: 2020
DOI: https://doi.org/10.48550/arXiv.2002.06261

Paper ID: 14
Authors: Rahul Singh, Karan Jindal, Yufei Yu, et al.
Title: Robustness tests of NLP machine learning models: Search and semantically replace
Venue: arXiv preprint
Details: arXiv:2104.09978
Year: 2021
DOI: https://doi.org/10.48550/arXiv.2104.09978

Paper ID: 15
Authors: Pin Ji, Yang Feng, Ruohao Zhang, et al.
Title: NLPLego: Assembling Test Generation for Natural Language Processing Applications
Venue: ACM Transactions on Software Engineering and Methodology
Year: 2023
DOI: https://doi.org/10.48550/arXiv.2302.10499

Paper ID: 16
Authors: Muhammad Hilmi Asyrofi, Zhou Yang, Imam Nur Bani Yusuf, et al.
Title: Biasfinder: Metamorphic test generation to uncover bias for sentiment analysis systems
Venue: IEEE Transactions on Software Engineering
Details: 48(12):5087–5101
Year: 2021
DOI: https://doi.org/10.1109/TSE.2021.3136169

Paper ID: 17
Authors: Adel Rahimi, Shaurya Jain
Title: Testing the effectiveness of saliency-based explainability in NLP using randomized survey-based experiments
Venue: arXiv preprint
Details: arXiv:2211.15351

Paper ID: 18
Authors: Minjie Ding, Mingang Chen, Wenjie Chen, et al.
Title: Data Generation, Testing and Evaluation of Chinese Natural Language Processing in the Cloud
Venue: 2022 IEEE 7th International Conference on Smart Cloud (Smart-Cloud)
Details: Pages 91–96

Paper ID: 19
Authors: Zeyu Sun, Zhenpeng Chen, Jie Zhang, Dan Hao
Title: Fairness Testing of Machine Translation Systems
Venue: ACM Transactions on Software Engineering and Methodology

Paper ID: 20
Authors: Mingyue Jiang, Houzhen Bao, Kaiyi Tu, et al.
Title: Evaluating Natural Language Inference Models: A Metamorphic Testing Approach
Venue: 2021 IEEE 32nd International Symposium on Software Reliability Engineering (ISSRE)
Details: Pages 220–230

Paper ID: 21
Authors: Ester Hlavnova and Sebastian Ruder
Title: Empowering Cross-lingual Behavioral Testing of NLP Models with Typological Features
Venue: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)
Pages: 7181–7198

Paper ID: 22
Authors: Aditi Mishra, et al.

Title: PromptAid: Prompt Exploration, Perturbation, Testing and Iteration using Visual Analytics for Large Language Models
Venue: arXiv preprint
Link: `https://doi.org/10.48550/arXiv.2304.01964`
Year: 2023

Here is the Overleaf format for IDs 23–30:

Surveyed Papers

Paper ID: 23
Authors: Xuezhi Wang, Haohan Wang, Diyi Yang
Title: Measure and Improve Robustness in NLP Models: A Survey
Venue: arXiv preprint
Year: 2021
DOI: https://doi.org/10.48550/arXiv.2112.08313

Paper ID: 24
Authors: Arshdeep Sekhon, Yangfeng Ji, Matthew B. Dwyer, Yanjun Qi
Title: White-box Testing of NLP Models with Mask Neuron Coverage
Venue: arXiv preprint
Year: 2022
DOI: https://doi.org/10.48550/arXiv.2205.05050

Paper ID: 25
Authors: Jay DeYoung, Sarthak Jain, Nazneen Fatema Rajani, Eric Lehman, Caiming Xiong, Richard Socher, Byron C. Wallace
Title: ERASER: A Benchmark to Evaluate Rationalized NLP Models
Venue: arXiv preprint
Year: 2019
DOI: https://doi.org/10.48550/arXiv.1911.03429

Paper ID: 26
Authors: Ryan Steed, Swetasudha Panda, Ari Kobren, Michael Wick
Title: Upstream Mitigation is Not All You Need: Testing the Bias Transfer Hypothesis in Pre-trained Language Models
Venue: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)
Pages: 3524–3542
Year: 2022
DOI: https://doi.org/10.18653/v1/2022.acl-long.247

Paper ID: 27
Authors: Milad Moradi, Matthias Samwald
Title: Evaluating the Robustness of Neural Language Models to Input Perturbations
Venue: arXiv preprint
Year: 2021
DOI: https://doi.org/10.48550/arXiv.2108.12237

Paper ID: 28
Authors: J. D. Zamfirescu-Pereira, Bjoern Hartmann, Qian Yang
Title: Conversation Regression Testing: A Design Technique for Prototyping Generalizable Prompt Strategies for Pre-trained Language Models
Venue: arXiv preprint
Year: 2023
DOI: https://doi.org/10.48550/arXiv.2302.03154

Paper ID: 29
Authors: L. Pandia, A. Ettinger
Title: Sorting Through the Noise: Testing Robustness of Information Processing in Pre-trained Language Models
Venue: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing
Year: 2021
DOI: https://doi.org/10.48550/arXiv.2109.12393

Paper ID: 30
Authors: Tal Golan, Matthew Siegelman, Nikolaus Kriegeskorte, Christopher Baldassano
Title: Testing the Limits of Natural Language Models for Predicting Human Language Judgements
Venue: Nature Machine Intelligence
Volume: 5
Number: 9
Pages: 952–964
Year: 2023
DOI: https://doi.org/10.1038/s42256-023-00718-1
Paper ID: 31
Authors: Frank Pallas, Dimitri Staufer, Jörn Kuhlenkamp
Title: Evaluating the Accuracy of Cloud NLP Services Using Ground-Truth Experiments
Venue: 2020 IEEE International Conference on Big Data (Big Data)

Pages: 341–350
Year: 2020
DOI: https://doi.org/10.1109/BigData50022.2020.9378188

Paper ID: 32
Authors: Qingqing Cao, Aruna Balasubramanian, Niranjan Balasubramanian
Title: Towards Accurate and Reliable Energy Measurement of NLP Models
Venue: arXiv preprint
Year: 2020
DOI: https://doi.org/10.48550/arXiv.2010.05248

Paper ID: 33
Authors: Dastan Hussen Maulud, Subhi RM Zeebaree, Karwan Jacksi, Mohammed A Mohammed Sadeeq, Karzan Hussein Sharif
Title: State of Art for Semantic Analysis of Natural Language Processing
Venue: Qubahan Academic Journal
Volume: 1
Number: 2
Pages: 21–28
Year: 2021
DOI: https://doi.org/10.48161/qaj.v1n2a44

Paper ID: 34
Authors: Phyllis Ang, Bhuwan Dhingra, Lisa Wu Wills
Title: Characterizing the Efficiency vs. Accuracy Trade-off for Long-Context NLP Models
Venue: arXiv preprint
Year: 2022
DOI: https://doi.org/10.48550/arXiv.2204.07288

Paper ID: 35
Authors: Beakcheol Jang, Myeonghwi Kim, Gaspard Harerimana, Sang-ug Kang, Jong Wook Kim
Title: Bi-LSTM Model to Increase Accuracy in Text Classification: Combining Word2vec CNN and Attention Mechanism
Venue: Applied Sciences
Volume: 10
Number: 17
Pages: 5841
Year: 2020
DOI: https://doi.org/10.3390/app10175841

Paper ID: 36
Authors: Amrit Nagarajan, Sanchari Sen, Jacob R. Stevens, Anand Raghunathan
Title: AxFormer: Accuracy-Driven Approximation of Transformers for Faster, Smaller, and More Accurate NLP Models
Venue: 2022 International Joint Conference on Neural Networks (IJCNN)
Pages: 1–8
Year: 2022
DOI: https://doi.org/10.1109/IJCNN55064.2022.9892797

Paper ID: 37
Authors: Tianlu Wang, Rohit Sridhar, Diyi Yang, Xuezhi Wang
Title: Identifying and Mitigating Spurious Correlations for Improving Robustness in NLP Models
Venue: arXiv preprint
Year: 2021
DOI: https://doi.org/10.48550/arXiv.2110.07736

Paper ID: 38
Authors: Bairu Hou, Jinghan Jia, Yihua Zhang, Guanhua Zhang, Yang Zhang, Sijia Liu, Shiyu Chang
Title: TextGrad: Advancing Robustness Evaluation in NLP by Gradient-Driven Optimization
Venue: arXiv preprint
Year: 2022
DOI: https://doi.org/10.48550/arXiv.2212.09254

Paper ID: 39
Authors: Vyas Raina, Adian Liusie, Mark Gales
Title: Is LLM-as-a-Judge Robust? Investigating Universal Adversarial Attacks on Zero-shot LLM Assessment
Venue: arXiv preprint
Year: 2024
DOI: https://doi.org/10.48550/arXiv.2402.14016

Paper ID: 40
Authors: Pak Yuen Patrick Chan, Jacky Keung, Zhen Yang
Title: Validating Plagiarism Detection Systems with Metamorphic Testing
Venue: 2023 International Symposium on Educational Technology (ISET)
Pages: 132–137
Year: 2023
DOI: https://doi.org/10.1109/ISET58841.2023.00034

Paper ID: 41
Authors: Shun-Hang Li, Gang Zhou, Zhi-Bo Li, Ji-Cang Lu, Ning-Bo Huang
Title: The Causal Reasoning Ability of Open Large Language Models: A Comprehensive and Exemplary Functional Testing
Venue: 2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security (QRS)
Pages: 240–249
Year: 2023
DOI: https://doi.org/10.1109/QRS60937.2023.00032

Paper ID: 42
Authors: Pietro Liguori, Cristina Improta, Roberto Natella, Bojan Cukic, Domenico Cotroneo
Title: Who Evaluates the Evaluators? On Automatic Metrics for Assessing AI-Based Offensive Code Generators
Venue: Expert Systems with Applications
Volume: 225
Pages: 120073
Year: 2023
DOI: https://doi.org/10.1016/j.eswa.2023.120073

Paper ID: 43
Authors: Michael Evans, Dominik Soós, Ethan Landers, Jian Wu
Title: MSVEC: A Multidomain Testing Dataset for Scientific Claim Verification
Venue: Proceedings of the Twenty-fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing
Pages: 504–509
Year: 2023
DOI: https://doi.org/10.1145/3565287.3617630

Paper ID: 44
Authors: Daniel Rodriguez-Cardenas, David N. Palacio, Dipin Khati, Henry Burke, Denys Poshyvanyk
Title: Benchmarking Causal Study to Interpret Large Language Models for Source Code
Venue: 2023 IEEE International Conference on Software Maintenance and Evolution (ICSME)
Pages: 329–334

Paper ID: 45
Authors: Mingxuan Xiao, Yan Xiao, Hai Dong, Shunhui Ji, Pengcheng Zhang
Title: LEAP: Efficient and Automated Test Method for NLP Software
Venue: 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE)

Paper ID: 46
Authors: Jingyue Li, Per Håkon Meland, Jakob Svennevik Notland, André Storhaug, Jostein Hjortland Tysse
Title: Evaluating the Impact of ChatGPT on Exercises of a Software Security Course
Venue: 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)

Paper ID: 47
Authors: Sergio Morales, Robert Clarisó, Jordi Cabot
Title: Automating Bias Testing of LLMs
Venue: 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE)

Paper ID: 48
Authors: Boxi Yu, Yiyan Hu, Qiuyang Mang, Wenhan Hu, Pinjia He
Title: Automated Testing and Improvement of Named Entity Recognition Systems
Venue: Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering

Paper ID: 49
Authors: Caroline Craig, Kartik Goyal, Gregory Crane, Farnoosh Shamsian, David A. Smith
Title: Testing the Limits of Neural Sentence Alignment Models on Classical Greek and Latin Texts and Translations
Venue: CEUR Workshop Proceedings
Volume: 3558
Pages: 73
Year: 2023
Link: https://ceur-ws.org/Vol-3558/paper6193.pdf

Paper ID: 50
Authors: Xiaoyuan Xie, Shuo Jin, Songqiang Chen
Title: qaAskeR+: A Novel Testing Method for Question Answering Software via Asking Recursive Questions
Venue: Automated Software Engineering
Volume: 30
Number: 1
Pages: 14
Year: 2023
DOI: https://doi.org/10.1145/1376616.1376746

Paper ID: 51
Authors: Elvys Soares, Manoel Aranda, Naelson Oliveira, Márcio Ribeiro, Rohit Gheyi, Emerson Souza
Title: Manual Tests Do Smell! Cataloging and Identifying Natural Language Test Smells
Venue: 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)
Year: 2023
DOI: https://doi.org/10.1109/ESEM56168.2023.10304800

Paper ID: 52
Authors: Dimitrios P. Panagoulias, et al.
Title: Evaluating the Potential of LLMs and ChatGPT on Medical Diagnosis and Treatment
Venue: 2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA)
Year: 2023
DOI: https://doi.org/10.1109/IISA59645.2023.10345968

Paper ID: 53

Authors: Jun Wang, et al.
Title: Back Deduction Based Testing for Word Sense Disambiguation Ability of Machine Translation Systems
Venue: Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis
Year: 2023
DOI: https://doi.org/10.1145/3597926.359808

Paper ID: 54
Authors: Christoph Treude and Hideaki Hata
Title: She Elicits Requirements and He Tests: Software Engineering Gender Bias in Large Language Models
Venue: 2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR)
Year: 2023
DOI: https://doi.org/10.1109/MSR59073.2023.00088

Paper ID: 55
Authors: Ezekiel Soremekun, Sakshi Udeshi, Sudipta Chattopadhyay
Title: Astraea: Grammar-based Fairness Testing
Venue: IEEE Transactions on Software Engineering
Year: 2022
Pages: 5188–5211
DOI: https://doi.org/10.1109/TSE.2022.3141758

Paper ID: 56
Authors: Valerio Basile, et al.
Title: UINAUIL: A Unified Benchmark for Italian Natural Language Understanding
Venue: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)
Year: 2023
DOI: https://doi.org/10.18653/v1/2023.acl-demo.33

Paper ID: 57
Authors: Qingchao Shen, et al.
Title: Natural Test Generation for Precise Testing of Question Answering Software
Venue: Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering
Year: 2022
DOI: https://doi.org/10.1145/3551349.3556953

Paper ID: 58
Authors: Yunhe Feng, et al.
Title: Investigating Code Generation Performance of ChatGPT with Crowd-sourcing Social Data
Venue: 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)
Year: 2023
DOI: https://doi.org/10.1109/COMPSAC57700.2023.00117

Paper ID: 59
Authors: Zhiyuan Yu, et al.
Title: CODEIPPROMPT: Intellectual Property Infringement Assessment of Code Language Models
Venue: International Conference on Machine Learning (PMLR)
Year: 2023
DOI: https://proceedings.mlr.press/v202/yu23g.html

Paper ID: 60
Authors: Sajed Jalil, et al.
Title: ChatGPT and Software Testing Education: Promises & Perils
Venue: 2023 IEEE International Conference on Software Testing, Verification, and Validation Workshops (ICSTW)
Year: 2023
DOI: https://doi.org/10.1109/ICSTW58534.2023.00078

Paper ID: 61
Authors: Wenxuan Wang, et al.
Title: MTTM: Metamorphic Testing for Textual Content Moderation Software
Venue: 2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)
Year: 2023
DOI: https://doi.org/10.1109/ICSE48619.2023.00200

Paper ID: 62
Authors: Jhon Bryan J. Cantil and Kristine Mae M. Adlaon
Title: Technology-Enhanced Systemic Quality Assurance with the Aid of Text-Based Emotion Recognition of Facebook Comments for Higher Education Institutions
Venue: AICCC '22: Proceedings of the 2022 5th Artificial Intelligence and Cloud Computing Conference
Year: 2022

DOI: `https://doi.org/10.1145/3582099.35821`

Paper ID: 63
Authors: Wentao Gao, Jiayuan He, and Van-Thuan Pham
Title: Metamorphic Testing of Machine Translation Models Using Back Translation
Venue: 2023 IEEE/ACM International Workshop on Deep Learning for Testing and Testing for Deep Learning (DeepTest)
Year: 2023
DOI: `https://doi.org/10.1109/DeepTest59248.2023.00008`

Paper ID: 64
Authors: Simin Chen, et al.
Title: NMTSLOTH: Understanding and Testing Efficiency Degradation of Neural Machine Translation Systems
Venue: ESEC/FSE 2022: Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering
Year: 2022
DOI: `https://doi.org/10.1145/3540250.354910`

Paper ID: 65
Authors: Shailja Thakur, et al.
Title: Benchmarking Large Language Models for Automated Verilog RTL Code Generation
Venue: 2023 Design, Automation & Test in Europe Conference & Exhibition (DATE)
Year: 2023
DOI: `https://doi.org/10.23919/DATE56975.2023.10137086`

Paper ID: 66
Authors: Ansong Ni, et al.
Title: Lever: Learning to Verify Language-to-Code Generation with Execution
Venue: International Conference on Machine Learning (PMLR)
Year: 2023
DOI: `https://proceedings.mlr.press/v202/ni23b.html`

Paper ID: 67
Authors: Binggang Zhuo, Masaki Murata, and Qing Ma
Title: Auxiliary Loss for BERT-Based Paragraph Segmentation
Venue: IEICE TRANSACTIONS on Information and Systems

Paper ID: 68
Authors: Debora Nozza, Federcio Bianchi, and Dirk Hovy
Title: Pipelines for Social Bias Testing of Large Language Models
Venue: Proceedings of BigScience Episode# 5–Workshop on Challenges & Perspectives in Creating Large Language Models
Year: 2022
DOI: https://doi.org/10.48550/arXiv.2005.04118

Paper ID: 69
Authors: Yuechen Li, et al.
Title: A Distance-Based Dynamic Random Testing Strategy for Natural Language Processing DNN Models
Venue: 2022 IEEE 22nd International Conference on Software Quality, Reliability and Security (QRS)
Year: 2022
DOI: https://doi.org/10.1109/QRS57517.2022.00089

Paper ID: 70
Authors: Marco Tulio Ribeiro, et al.
Title: Beyond Accuracy: Behavioral Testing of NLP Models with CheckList
Venue: arXiv preprint
Year: 2020
DOI: https://doi.org/10.48550/arXiv.2005.04118

Paper ID: 71
Authors: Andrey Zvonarev and Andrey Bilyi
Title: A Comparison of Machine Learning Methods of Sentiment Analysis Based on Russian Language Twitter Data
Venue: MICSECS 2019
Year: 2019
Link:

Paper ID: 72
Authors: Pingchuan Ma, Shuai Wang, and Jin Liu
Title: Metamorphic Testing and Certified Mitigation of Fairness Violations in NLP Models
Venue: IJCAI
Year: 2020
Link: https://www.ijcai.org/proceedings/2020/0064.pdf

Paper ID: 73
Authors: Najlah Gali, et al.
Title: Framework for Syntactic String Similarity Measures
Venue: Expert Systems with Applications
Year: 2019
DOI: https://doi.org/10.1016/j.eswa.2019.03.048

Paper ID: 74
Authors: Liqun Sun and Zhi Quan Zhou
Title: Metamorphic Testing for Machine Translations: MT4MT
Venue: 2018 25th Australasian Software Engineering Conference (ASWEC)
Year: 2018
DOI: https://doi.org/10.1109/ASWEC.2018.00017

Paper ID: 75
Authors: Robert White and Jens Krinke
Title: TestNMT: Function-to-Test Neural Machine Translation
Venue: NL4SE 2018: Proceedings of the 4th ACM SIGSOFT International Workshop on NLP for Software Engineering
Year: 2018
DOI: https://doi.org/10.1145/3283812.3283823

Paper ID: 76
Authors: Pooja Malik and Anurag Singh Baghel
Title: Performance Enhancement of Machine Translation Evaluation Systems for English–Hindi Language Pair
Venue: International Journal of Modern Education and Computer Science
Year: 2019
DOI: https://doi.org/10.5815/ijmecs.2019.02.06

Paper ID: 77
Authors: Tao Yu, et al.
Title: Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task
Venue: arXiv preprint
Year: 2018
DOI: https://doi.org/10.48550/arXiv.1809.08887

Paper ID: 78
Authors: Bassam Hammo, Martha W. Evens, and Hani Abu-Salem
Title: Effects of the College Major on Assessments of Arabic Text Summaries