**Deep Learning Mini-Project Datasets**

*English (Base Language)*

Clarification of coloring scheme used in the subsequent table:

* Blue: Training Set
* Yellow: Benchmarking Set
* Green: Adversarial Training

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Link | Size | Task Description |
| Learning From the Worst (Dynamically generated hate speech dataset) | [Publication](https://aclanthology.org/2021.acl-long.132/)  [Dataset](https://github.com/bvidgen/Dynamically-Generated-Hate-Speech-Dataset) | 41’225 | Multi-category hate speech detection |
| Overview of the Task on Automatic Misogyny Identification at IberEval 2018 | [Publication](http://ceur-ws.org/Vol-2150/overview-AMI.pdf)  [Dataset](https://amiibereval2018.wordpress.com/important-dates/data/) | 3’977 | Binary (misogyny / not), 5 categories (stereotype, dominance, derailing, sexual harassment, discredit), target of misogyny (active or passive) |
| CONAN – Counter Narratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech | [Publication](https://www.aclweb.org/anthology/P19-1271.pdf)  [Dataset](https://github.com/marcoguerini/CONAN) | 1’288 (possibly more in total) | Binary (Islamophobic / not), multi-topic (Culture, Economics, Crimes, Rapism, Terrorism, Women Oppression, History, Other/generic) |
| Multilingual and Multi-Aspect Hate Speech Analysis | [Publication](https://arxiv.org/abs/1908.11049)  [Dataset](https://github.com/HKUST-KnowComp/MLMA_hate_speech) | 5’647 | Detailed taxonomy with cross-cutting attributes: Hostility, Directness, Target attribute and Target group. |
| ToxiGen: A Large-Scale Machine-Generated Dataset for Adversarial and Implicit Hate Speech Detection | [Publication](https://arxiv.org/pdf/2203.09509.pdf)  [Dataset](https://github.com/microsoft/ToxiGen) | - | Adversarial Training on Hate Speech examples. |

*French (Training Data for Finetuning)*

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Link | Size | Task Description |
| CONAN – Counter Narratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech | [Publication](https://www.aclweb.org/anthology/P19-1271.pdf)  [Dataset](https://github.com/marcoguerini/CONAN) | 1’719 (possibly more in total) | Binary (Islamophobic / not), multi-topic (Culture, Economics, Crimes, Rapism, Terrorism, Women Oppression, History, Other/generic) |
| Multilingual and Multi-Aspect Hate Speech Analysis | [Publication](https://arxiv.org/abs/1908.11049)  [Dataset](https://github.com/HKUST-KnowComp/MLMA_hate_speech) | 4’014 | Detailed taxonomy with cross-cutting attributes: Hostility, Directness, Target attribute and Target group. |

*Arabic (Training Data for Finetuning)*

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Link | Size | Task Description |
| Multilingual and Multi-Aspect Hate Speech Analysis | [Publication](https://arxiv.org/abs/1908.11049)  [Dataset](https://github.com/HKUST-KnowComp/MLMA_hate_speech) | 3’353 | Detailed taxonomy with cross-cutting attributes: Hostility, Directness, Target attribute and Target group. |
| L-HSAB: A Levantine Twitter Dataset for Hate Speech and Abusive Language | [Publication](https://www.aclweb.org/anthology/W19-3512)  [Dataset](https://github.com/Hala-Mulki/L-HSAB-First-Arabic-Levantine-HateSpeech-Dataset) | 5’846 | Ternary (Hate, Abusive, Normal) |

*German (Training Data for Finetuning)*

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Link | Size | Task Description |
| RP-Mod & RP-Crowd: Moderator- and Crowd-Annotated German News Comment Datasets | [Publication](https://datasets-benchmarks-proceedings.neurips.cc/paper/2021/file/c9e1074f5b3f9fc8ea15d152add07294-Paper-round2.pdf)  [Dataset](https://zenodo.org/record/5291339#.Ybr_9VkxkUE) | 85’000 | Binary (Offensive or Not), Multi-class/-label (sexism, racism, threats, insults, profane language, meta, advertisement) |