

English bias in computing: Images to the rescue

September 6 2022



Overview of the tasks in IGLUE, which include grounded natural language inference, visual question answering, grounded reasoning, and cross-modal retrieval. Each task is associated with an example of input and output (English translations at the bottom). Credit: *Proceedings of The Thirty-ninth International Conference on Machine Learning* (2022). DOI: 10.48550/arXiv.2201.11732

So many languages; and yet English is allowed almost total domination when it comes to the AI technology Machine Learning (ML). If, for example, researchers are training a computer in grasping the content of a random text, the training samples will typically be in English.

"This introduces a significant unintentional cultural [bias](#). Even after extensive training, the machine will never have been exposed to bull taming in India, to Chinese hot pot cooking, or to other phenomena which are familiar to millions of people, but just happen to lie outside the native English-speaking horizon," says Ph.D. researcher Emanuele

Bugliarello, Department of Computer Science (DIKU), University of Copenhagen.

In a truly intercultural effort, Bugliarello and colleagues from a range of countries have created a new tool which encourages a more diverse approach. IGLUE (Image-Grounded Language Understanding Evaluation), as they have named the tool, is a benchmark which allows for scoring the efficiency of an ML solution in 20 languages (rather than English alone).

Their scientific article introducing IGLUE has been accepted for publication in the upcoming *Proceedings of The Thirty-ninth International Conference on Machine Learning*, one of the top conferences in the field.

Volunteers provided culture specific images

How can a new benchmark make a difference?

"When ML research teams create new solutions, they are always highly competitive. If another group has succeeded in solving a given ML task with 98 percent accuracy, you will try to get 99 percent and so forth. This is what drives the field forward. But the downside is that if you don't have a proper benchmark for a given feature, it will not be prioritized. This has been the case for multimodal ML, and IGLUE is our attempt to change the scene," says Bugliarello.

Basing training on images is standard in ML. However, the images are usually "labeled", meaning bits of text will accompany each image, aiding the learning process of the machine. While the labels are normally in English, IGLUE covers 20 typologically diverse languages, spanning 11 language families, 9 scripts, and 3 geographical macro-areas.

A part of the images in IGLUE are culture specific. These images were obtained through a mail campaign. The researchers asked volunteers in geographically diverse countries to provide images and texts in their natural language and preferably about things that were important in that country.

Overwhelmed by positive reactions

The current lack of multimodal ML does have practical implications, Bugliarello explains:

"Let us say you have a food allergy, and you have an app which can tell you if the problematic ingredients are present in a meal. Finding yourself at a restaurant in China, you realize the menu is all in Chinese but has pictures. If your app is good, it may translate the picture into a recipe—but only if the machine was exposed to Chinese samples during training."

In other words, non-English speakers get a poorer version of ML-based solutions:

"The performance of many top ML solutions will drop instantly, as they become exposed to data from non-English speaking countries. And notably, the ML solutions miss out on concepts and ideas that are not formed in Europe or North America. This is something which the ML research community needs to address," Bugliarello says.

Fortunately, many colleagues have seen the light, Bugliarello notes:

"This all began a few years ago when we wrote a paper for the EMNLP conference (Empirical Methods in Natural Language Processing). We just wanted to point to an issue, but were soon overwhelmed with interest, and much to our surprise our contribution was selected as Best

Long Paper. People clearly saw the problem, and we were encouraged to do more."

May help the visually impaired

Sometimes the current success almost feels like a burden, Bugliarello admits:

"As a public university, we do have limited resources. We cannot pursue all aspects of this huge task. Still, we can see that other groups are joining in. We can also feel interest from the large tech corporations. They are strongly engaged in ML and are beginning to realize how English bias can be a problem. Obviously, they are not happy to see the performance of their solutions dropping significantly when applied outside English-language contexts."

Despite the positive developments, Bugliarello does not allow himself to get carried away. Asked how close we are to achieving non-biased Machine Learning, he answers:

"Oh, we are very far away."

Still, this is not only about cultural equality:

"The methodology behind IGLUE may find several applications. For example, we hope to improve solutions for visually impaired. Tools exist, which help visually impaired in following the plot of a movie or another type of visual communication. These tools are currently far from perfect, and I would very much like to be able to improve them. This is a bit further into the future, though," Bugliarello says

More information: Emanuele Bugliarello et al, IGLUE: A Benchmark for Transfer Learning across Modalities, Tasks, and Languages,

Proceedings of The Thirty-ninth International Conference on Machine Learning (2022). [DOI: 10.48550/arXiv.2201.11732](https://doi.org/10.48550/arXiv.2201.11732)

Provided by University of Copenhagen

Citation: English bias in computing: Images to the rescue (2022, September 6) retrieved 26 April 2024 from <https://techxplore.com/news/2022-09-english-bias-images.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.