Introduction to AIRankings

Overview

AIRankings (https://airankings.org/) is a metric-based ranking that provides insights into Artificial Intelligence research capability worldwide at various levels (e.g., authors, institutes, cities, countries). There are two indicators in the AIRankings to operationalize the AI research capability: Adjusted Publications and AI Index. While Adjusted Publications is based on the number of publications at top AI venues in core AI areas, AI Index adopts a more holistic and multidisciplinary view of AI research by calculating the geometric average of Adjusted Publications for selected areas.

Why do we need a new ranking?

In the new era of big innovation, AI is an emerging field of importance that has been separated from the traditional fields such as computer science and statistics. It has been treated as an independent discipline nowadays when many academic institutes establish their own AI departments and educational programs.

The development of AI research witnessed several ups and downs after its birth in 1956 when researchers started to ask about the possibility of human-like machines to simulate aspects of intelligence. Though the development of AI has been powered by tremendous progress in specialized intelligence systems, the big question of general artificial intelligence starts to regain its importance in the 21st century.

Rankings are valuable as a reference for decision-making and as an incentive for improvement. Though many institutes around the world launch their own ranking systems, AIRankings is among the very few that mainly focus on the field of Artificial Intelligence. Besides, AIRankings echoes the idea of general artificial intelligence by the inclusion of AI Index that encourages multidisciplinary research.

AIRankings has the following characteristics:

- **Objectivity**: Unlike rankings such as those published by US News & World Report, QS, or THE that gather information via surveys, AIRankings is a transparent metric-based system that minimizes the influence of subjective opinions by adopting a mathematical algorithm and publicly available data.
- **Comprehensiveness**: CSRankings is another metric-based ranking of top computer science institutions, which also has an AI subfield. The differences between AIRankings and CSRankings lie in that 1) AIRankings mainly focuses on AI research, leading to a more detailed categorization of the core areas within AI research; and 2) AIRankings has a broader coverage of AI-relevant conferences and journals while CSRankings only includes publications from conferences. While CSRankings is useful for those who want to gain insights into the broader field of computer science, AIRankings provides a closer and more in-depth understanding of Artificial Intelligence.
- **Multidisciplinary Approach**: AI is becoming influential by incorporating multiple subjects into a whole and is multidisciplinary by its nature. AIRankings recognizes the importance of

multidisciplinary research and creates AI Index to more accurately depict the AI research capability across multiple AI core areas. For the same number of publications, an author who has a more diverse research portfolio will receive a higher AI Index score.

- **Multi-level Dashboard:** Unlike other rankings that focus either on the institute-level or the author-level, AIRankings has a dashboard that presents results at multiple levels including authors, institutes, cities and, countries. It can generate insights from both the micro-level and the macro-level. Besides, AIRankings provides interactive visualization, presenting the time trajectory of AI research development within each level.

Methodology

The detailed methodology of AIRankings is elaborated below.

- **Core Areas:** AIRankings includes the following six core areas of AI: Computer Vision, Natural Language, Machine Learning, Cognitive Reasoning, Robotics, and Multi-Agent Systems. In addition, we also take into account two areas named Simulation for training AI agents in graphics environments like VR and AR, and AI-General for pursuing general AI research, including venues that encompass multiple areas of AI. These areas play in synergy and have been integrated towards building general AI systems. The corresponding top academic conferences/journals are fully included in each area. For example, the computer vision field includes CVPR, ICCV, ECCV, PAMI, IJCV, CVIU, PR, FTCGV; the natural language processing field venues such as ACL, EMNLP, COLING, CONLL, NAACL; the machine learning field JMLR, NECO, IJON, NeurIPS, ICML, COLT, AISTATS, ICLR, etc.
- **Author-Level:** Given a selected timespan, an author is the smallest unit of calculation and serves as the basis for the calculation of AI research capability at the other three levels.
    1. **Adjusted Publications** is the author's total publications in selected areas, adjusted by two factors: an article is weighted by the importance of its venue, and an article with K senior co-authors will give 1/K score to each senior co-author. Note that the venue weight has four levels, 0.25, 0.5, 0.75, and 1. Senior authors are faculty members who specialize in the AI field, which doesn't count students. In the formula below, n means the total number of an author's publication, P_i means the weight of the venue for the publication i, and k_i means the number of senior co-authors of publication i.

$$Author's\ Adjusted\ Publications\ Score = \sum_{i=1}^{n} \mathrm{P}_i * \frac{1}{k_i}$$

    2. **AI Index** is the geometric average of Author's Adjusted Publications Scores for the selected areas. When one area is selected, Adjusted Publications and AI Index are the same. When multiple areas are selected, the algorithm normalizes the sizes of areas. AI Index reflects the research breadth of an author, that is, under the same number of publications, an author will receive higher scores if they are published across multiple AI core areas than one single area. In the formula below, S_i means the Adjusted Publications score of an author in a selected area i, n means the number of selected areas.

$$Author's\ AI\ Index = \prod_{i=1}^{n} (S_i + 1)^{\frac{1}{n}} - 1$$

- **Institute-Level:** An institute's Adjusted Publications and AI Index are similar to those of an author but including the publications of all its senior authors. When an author changes affiliations, his/her historical scores move with him/herself to the new institute. This algorithm assumes that the AI research capability of an institute is determined by its current talents rather than its historical achievements.
    1. **Adjusted Publications** is the total number of publications of an institute in selected areas, which is the sum of all its authors' Adjusted Publications Scores. In the formula below, *m* means the total number of authors in the selected areas, *j* means the specific author *j*.

$$Institute's\ Adjusted\ Publications\ Score\ = \sum_{i=1}^{m}\sum_{j=1}^{n_i} \mathrm{P}_{ij} * \frac{1}{k_{ij}}$$

    2. **AI Index** is the geometric average of Institute's Adjusted Publications Scores in selected areas, which reflects the general AI research capability of an institute across AI core areas. In the formula below, $S_{I,i}$ means the Adjusted Publications score of an institute *I* in a selected area *i*, *N* means the number of selected areas.

$$Institute's\ AI\ Index = \prod_{i=1}^{N} (S_{I,i} + 1)^{\frac{1}{N}} - 1$$

- **City -Level & Country-Level:** AIRankings at the city-level and the country-level depicts the AI research capability by including all its senior authors in a certain city or a certain country, which follows the same algorithm as the institute-level indicators.

Shortcomings & Future Improvements

1. To focus on the professional AI research community, AIRankings includes active authors with Adjusted Publications ≥ 2.0 in all core areas in the past 10 years.
2. AIRankings currently collects data of research institutes in academia, which doesn't include those corporate-funded institutes.
3. Not all AI research results are published in AI-specific journals and conferences. Some papers are published in multidisciplinary science journals (e.g., Cell, Nature, Science), which are not included in the database.
4. AIRankings differentiates the weights of AI-relevant academic conferences and journals in the calculation based on their importance and the influence within the field. The weights are determined through the consultation process of AI experts.
5. The city-level here can be more accurately described as the metropolitan-level. For example, MIT is considered as part of the Boston area instead of Cambridge.