

MICROSOFT KDD CUP 2016 CONTEST

OFFICIAL RULES

PLEASE NOTE: It is your sole responsibility to comply with your employer's gift policies. If your participation violates your employer's policies, you may be disqualified. Microsoft disclaims any and all liability and will not be party to any disputes or actions related to this matter.

GOVERNMENT EMPLOYEES: Microsoft is committed to complying with government gift and ethics rules and therefore government employees are not eligible to receive any prizes, but may participate in the competition.

This Contest is hosted in the United States, and entry information is collected on computers in the United States. This Contest will be governed by the laws of the State of Washington, and you consent to the exclusive jurisdiction and venue of the courts of the State of Washington for any disputes arising out of this Contest. **If you do not agree with this provision and these Official Rules, please do not enter this Contest.**

COMMON TERMS USED IN THESE RULES

These are the official rules that govern how the Microsoft KDD Cup 2016 contest will operate ("Contest").

In these rules, "we," "our," and "us" refer to Microsoft Corporation, the sponsor of this Contest. "You" refers to an eligible Contest entrant.

CONTEST DESCRIPTION

The object of this contest is to create the best ranking prediction model that ranks the most influential research institutions in the world using publicly available information and datasets. For the purposes of this Contest, each ranked list of research institutions you submit in the Contest will be called an "entry." All eligible entries received will be evaluated according to the Judging Criteria described below to determine winners.

WHAT ARE THE START AND END DATES?

This Contest starts at 12:01 AM Pacific Time (PT) on February 27, 2016, and ends at 11:59 PM PT on August 13, 2016 and will consist of multiple rounds as follows:

Round 1: General Entry Period

This General Entry Period will consist of three (3) Phases. Entries must be submitted during each Phase to be eligible for scoring.

Phase 1: 12:01 AM PT, **February 27**, 2016 to 11:59 PM PT, **March 28**, 2016

Phase 2: 12:01 AM PT, **March 29**, 2016 to 11:59 PM PT, **April 28**, 2016

Phase 3: 12:01 AM PT, **April 29**, 2016 to 11:59 PM PT, **May 28**, 2016

Round 2: Top-Ranked Finalist Paper Submission Period

During this Period, selected top-ranked finalists will be asked to write a paper describing their entry. This Period will begin immediately after winner notification and will end at 11:59 PM PT on July 15, 2016.

Round 3: Top-Ranked Finalist KDD Cup Workshop

The KDD Cup Workshop will be held on August 13, 2016 and will consist of top-ranked finalists presenting their findings live at the event or via video conference.

CAN I ENTER?

Yes! The KDD Cup 2016 Contest is for researchers and practitioners from data science, data mining, knowledge discovery, large-scale data analytics, and big data. However, you are only eligible to receive the prizes in this Contest if you meet the following requirements at the time of your entry:

- **You are 14 years of age or older as of the start of the contest; and**
 - If you are 14 years of age or older, but are considered a minor in your place of residence, you must have your parent's or legal guardian's permission prior to submitting an entry into this Contest;
- **You do NOT** currently reside in any of the following countries: Cuba, Iran, North Korea, Sudan and Syria.
 - **PLEASE NOTE:** U.S. export regulations prohibit the export of goods and services to Cuba, Iran, North Korea, Sudan and Syria. Therefore, residents of these countries / regions are not eligible to participate.
- **You are NOT** an employee of Microsoft Corporation or an employee of a Microsoft subsidiary; **and**
- **You are NOT** involved in any part of the administration and execution of this Contest.

This Contest is void outside the geographic area described above and wherever else prohibited by law.

HOW DO I ENTER?

Round 1: General Entry Period

To participate, you must first:

- Go to <http://aka.ms/kddcup>, sign in, and register your team by providing:
 - a. The first and last name of your team leader. Enter your name if you are the team leader.
 - b. A unique name for your team. The team name will be used as a unique identifier and may only contain alphanumeric characters

NOTE REGARDING TEAMS: You are only allowed to be a member of one team. If one person participates in multiple teams, all teams associated with that individual will be disqualified. Teams must be fully formed and locked by 11:59 PM on Friday, May 20, 2016. No changes to teams will be allowed after this date.

Next, visit <http://aka.ms/kddcup> and follow the on-screen instructions to develop your entry, leveraging the Details for background, task details, and dataset information. You will only be required to develop a prediction model that speaks to the list of affiliation names and ID's we make available. The team leader must upload the final entry file to the website.

While there is no limit to the number of entries you can submit, we will only evaluate the last submission in each phase.

Round 2: Top-Ranked Finalist Paper Submission Period

During this round, selected top-ranked finalists will need to submit a paper following KDD Workshop format describing their solution in generating the entry. The team leader must submit their paper no later than the submission close date above.

Round 3: Top-Ranked Finalist KDD Cup Workshop

During this round, Top-Ranked finalists will be invited to present their entry and paper at the KDD Cup 2016 Workshop in San Francisco.

We are not responsible for entries that we do not receive for any reason, or for entries that we receive but are not decipherable for any reason.

We will automatically disqualify any incomplete or illegible entries.

WHAT CONSTITUTES AN ELIGIBLE ENTRY?

To be eligible for judging, an entry must meet the following content / technical requirements:

- Your entry must be a single .tsv text file that contains a row for each item in the following format:
 - [conference id] \t [affiliation id] \t [probability score] \n
 - The Conference ID and the **Affiliation ID** must match the format given in the graph data. The probability score is a decimal value between 0 and 1 representing the importance of an affiliation in relation to the given conference. If an affiliation is missing from the results, the value will be assumed as 0.
 - The filename must be **results.tsv**
- Teams are welcome to use external data as long as it is publicly available.

In addition:

- Your entry must be your own original work; **and**
- You must have obtained any and all consents, approvals or licenses required for you to submit your entry; **and**

Entries may **NOT** contain, as determined by us, in our sole and absolute discretion, any content that:

- Is sexually explicit, unnecessarily violent or derogatory of any ethnic, racial, gender, religious, professional or age group; profane or pornographic;
- Promotes alcohol, illegal drugs, tobacco, firearms/weapons (or the use of any of the foregoing) or a particular political agenda;
- Is obscene or offensive;
- Defames, misrepresents or contains disparaging remarks about other people or companies;
- Communicates messages or images inconsistent with the positive images and/or good will to which we wish to associate; and/or violates any law;

We reserve the right to reject any entry, in our sole and absolute discretion, that we determine does not meet the above criteria.

HOW WILL MY ENTRY BE POTENTIALLY USED?

Other than what is set forth below, we are not claiming any ownership rights to your entry. However, by submitting your entry, you:

- Are granting us an irrevocable, royalty-free, worldwide right and license to: (i) use, review, assess, test and otherwise analyze your entry and all its content in connection with this Contest; and (ii) feature your entry and all content in connection with the marketing, sale, or promotion of this Contest (including but not limited to internal and external sales meetings, conference presentations, tradeshow and screen shots of the Contest entry in press releases) in all media (now known or later developed);
- Agree to sign any necessary documentation that may be required for us and our designees to make use of the rights you granted above;
- Understand and acknowledge that the Sponsor(s) may have developed or commissioned materials similar or identical to your submission, and you waive any claims you may have resulting from any similarities to your entry;
- Understand that we cannot control the incoming information you will disclose to our representatives in the course of entering, or what our representatives will remember about your entry. You also understand that we will not restrict work assignments of representatives who have had access to your entry. By entering this Contest, you agree that use of information in our representatives' unaided memories in the development or deployment of our products or services does not create liability for us under this agreement or copyright or trade secret law;
- Understand that you will not receive any compensation or credit for use of your entry, other than what is described in these Official Rules

Please note that following the end of this Contest your entry may be posted on a website selected by us for viewing by visitors to that website. We are not responsible for any unauthorized use of your entry by visitors to this website. While we reserve these rights, we are not obligated to use your entry for any purpose, even if it has been selected as a winning entry.

If you do not want to grant us these rights to your entry, please do not enter this Contest.

WINNER DETERMINATION AND PRIZES

After the close of all three Phases, the top 3 winners will be determined by the judging criteria below.

JUDGING CRITERIA

Please refer to the Evaluation section of this website for more information.

The evaluation will be conducted in three phases. In each phase, we will only choose one conference to evaluate each team's results. A team's final result will be calculated as the weighted sum of the chosen three conferences' submissions, i.e., $20\% \times \text{Conf1} + 40\% \times \text{Conf2} + 40\% \times \text{Conf3}$.

The ground truth will be determined by all the full research papers accepted in each conference. We use the following simple policy to determine the Affiliation Ranking:

1. Each accepted paper has an equal vote (i.e., they are equally important).
2. Each author has an equal contribution to a paper.
3. If an author has multiple affiliations, each affiliation also contributes equally.

Once we obtain the above ground truth ranking or relevance score, we will utilize the metric NDCG@20 to measure the relevance.

In the submitted ranking results, we ask the participants to avoid generating the same ranking scores for any two affiliations. If we find the ranking scores are the same for some affiliations, we will rank these affiliations based on the order of them in the submitted results.

In the event of a tie between any eligible entries, we will evaluate the NDCG beyond the top 20 positions until we break the ties.

LEADERBOARD

In this challenge, we will also provide a real-time leaderboard to show the progress of each team. Since the ground-truth data are unknown until the conference acceptance results are revealed, we will use the aggregated consensus data from all the participating teams to rank each team in the leaderboard.

We will evaluate each team's success in each phase once we obtain the ground truth paper acceptance data, and will announce the results of each stage as soon as possible, so teams can utilize such information to adjust their models.

CONTEST PRIZES

The top three (3) ranked finalists who complete the remaining entry requirements will be declared official winners on August 13, 2016.

One (1) Grand Prize. A check for \$10,000 USD divided equally among each team member.

One (1) First Prize(s). A check for \$6,500 USD divided equally among each team member.

One (1) Second Prize(s). A check for \$3,500 USD divided equally among each team member.

Limit one (1) prize per person per person and one prize overall per team.

*NOTE: For residents of Brazil exclusively, you will be awarded a gift card from a local merchant in the amount of your cash prize.

If you are a potential winner, we will notify you by sending a message to the email address, the phone number, or mailing address (if any) provided at time of entry within seven (7) days following completion of judging. If the notification that we send is returned as undeliverable, or you are otherwise unreachable for any reason, we may award to a runner-up.

If there is a dispute about who the potential winner is, we will consider the potential winner to be the authorized account holder of the email address used to enter the Contest. If you are a potential winner, we may require you

to sign an Affidavit of Eligibility, Liability/Publicity Release and a W-9 tax form or W-8 BEN tax form within 10 days of notification. If you complete a tax form, you will be issued an IRS Form 1099 the following January, for the actual value of the prize. You are advised to seek independent counsel regarding the tax implications of accepting a prize. If you are a potential winner and you are 14 or older, but are considered a minor in your place of legal residence, we may require your parent or legal guardian to sign all required forms on your behalf. If you do not complete the required forms as instructed and/or return the required forms within the time period listed on the winner notification message, we may disqualify you and select a runner-up.

If you are confirmed as a winner of this Contest:

- You may not exchange your prize for cash or any other merchandise or services. However, if for any reason an advertised prize is unavailable, we reserve the right to substitute a prize of equal or greater value; **and**
- You may not designate someone else as the winner. If you are unable or unwilling to accept your prize, we may award it to a runner up; **and**
- If you accept a prize, you will be solely responsible for all applicable federal, state, and local taxes related to accepting the prize; **and**
- If you are otherwise eligible for this Contest, but are considered a minor in your place of residence, we may award the prize to your parent/legal guardian on your behalf; **and**
- Unless otherwise noted, all prizes are subject to their manufacturer's warranty and / or terms and conditions.

WHAT OTHER CONDITIONS AM I AGREEING TO BY ENTERING?

By entering this Contest, you agree:

- To abide by these Official Rules; **and**
- To release and hold harmless Microsoft and its respective parents, subsidiaries, affiliates, employees and agents from any and all liability or any injury, loss or damage of any kind arising from or in connection with this Contest, or any prize won; **and**
- That Microsoft's decisions will be final and binding on all matters related to this Contest; **and**
- That, by accepting a prize, Microsoft may use your proper name and state of residence online and in print, or in any other media, in connection with this Contest, without payment or compensation to you, except where prohibited by law.

WHAT LAWS GOVERN THE WAY THIS CONTEST IS EXECUTED AND ADMINISTRATED?

This Contest will be governed by the laws of the State of Washington, and you consent to the exclusive jurisdiction and venue of the courts of the State of Washington for any disputes arising out of this Contest.

WHAT IF SOMETHING UNEXPECTED HAPPENS AND THE CONTEST CAN'T RUN AS PLANNED?

If someone cheats, or a virus, bug, bot, catastrophic event or any other unforeseen or unexpected event that cannot be reasonably anticipated or controlled (also referred to as force majeure), affects the fairness and / or integrity of this Contest, we reserve the right to cancel, change or suspend this Contest. This right is reserved whether the event is due to human or technical error. If a solution cannot be found to restore the integrity of the Contest, we reserve the right to select winners from among all eligible entries received before we had to cancel, change or suspend the Contest.

If you attempt or we have strong reason to believe that you have compromised the integrity or the legitimate operation of this Contest by cheating, hacking, creating a bot or other automated program, or by committing fraud in ANY way, we may seek damages from you to the fullest extent permitted by law. Further, we may disqualify you and ban you from participating in any of our future Contests, so please play fairly.

HOW CAN I FIND OUT WHO WON?

We will post the names of winners who received a prize online at <http://aka.ms/kddcup>. This list will remain posted one month after winner selection.

WHO IS SPONSORING THIS CONTEST?

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052