Registration “Landing Page”

# $50,000 BANNER Challenge

The millions of biomedical publications that exist are a valuable, but difficult to handle resource. Identifying those documents that are most relevant to particular diseases is currently a costly, human intensive activity. The Scripps Research Institute, together with the NASA Tournament Lab at Harvard, is launching a contest to develop new algorithms to aid in the automated Named Entity Recognition (NER) of biomedical publications.

The goal of this contest is to develop an algorithm to generate keywords based on the occurrence of specific entities such as genes, proteins, diseases, as well as the analysis of natural language text of publication abstracts. A large set of manually annotated abstracts has been created to aid in algorithm development and scoring. A full problem description will be provided only after the registration period.

To participate, entrants are required to register for the contest between [DATE REGISTRATION START] and [DATA REGISTRATION END].

# Contest Format

Registration is limited to 350 Top Coder members. Only members with some experience in at least one Top Coder competition (either MM or ALGO) are eligible to participate.

This is an experimental contest format and thus will be an unrated event.

A total of $50,000 in prizes will be distributed to competitors in the following manner.

$35,000 will be distributed in the form of room prizes to the winners in virtual competition rooms of no more than 14 participants. Additional $15,000 will be distributed in the form of grand prizes to the best performers across rooms.

The exact rules to determine the winners can differ across rooms. Some rooms will have a competition similar to a FirstToFinish contest and others will have a more traditional MM competition. At the end of the registration phase, we will send an email to all participants with an accurate description of the competition rules.

As part of the experimental contest format, participating in this contest also involves participating in a research project. Your participation is voluntary. If you choose to participate, you will need to complete two surveys: a registration survey and a final survey. You may discontinue your participation at any time.

The data collected, including survey responses, analysis of code submissions, and communications on the public forum of the challenge, will be used for research purposes. The substance of your survey responses will not affect your eligibility for winning a prize in this or future contests.

Only the immediate project team at TopCoder and researchers at Harvard University will see your individual data. Data will only be shared in an anonymous form in which individuals cannot be identified.

If you have any questions about the use of your information, please contact: [CONTACT PERSON]. Harvard University has a Standing Committee on the Use of Human Subjects in Research (CUHS) to which complaints or problems concerning any research project may, and should, be reported if they arise.

[AGREE TO REGISTER]

Communications After Registration

# Competition Rules for the Race

Dear [HANDLE],

Thanks for your interest in this competition. You have been randomly assigned to a room.

**In this room, the goal of the competition is to be the first to achieve a final score of at least [SCORE THRESHOLD].**

The first 2 coders to achieve a score greater than [SCORE THRESHOLD] will be awarded room prizes of $1,000 and $400 to the first and the second respectively. No prizes will be awarded if no submission reaches a score of at least [SCORE THRESHOLD].

You will be also competing for 3 additional grand prizes across all rooms with the same competition style. Such grand prizes are distributed as follows:

1. $3,000 to be the first to achieve at least [SCORE THRESHOLD]
2. $1,500 to be the second to achieve at least [SCORE THRESHOLD]
3. $500 to be the third to achieve at least [SCORE THRESHOLD].

So you can win up to $4,000.

At each submission, Top Coder will compute the final score in addition to the partial score. Final scores are not going to be shown on the leaderboard which displays the usual things (partial scores, temporary ranking, etc.).

As soon as a competitor in your room achieves a final score of at least [SCORE THRESHOLD], Top Coder will send an email to all room members to communicate that the first room prize has been awarded, but the second room prize is still available. The same thing happens when also a second competitor achieves a final score of at least [SCORE THRESHOLD], and then the competition is over.

Note, you are not allowed to discuss any of the details of your competition room with other members of the community during the 2 weeks of the competition.

For any question please contact [PERSON CONTACT].

# Competition Rules for the Tournament

Dear [HANDLE],

Thanks for your interest in this competition. You have been randomly assigned to a room.

**In this room, the goal of the competition is to achieve the best final score.**

The top 2 coders to achieve the highest final scores will be awarded room prizes of $1,000 and $400 to the first and the second respectively.

You will be also competing for additional grand prizes across all rooms with the same competition style. Such grand prizes are distributed as follows:

1. $3,000 to achieve the highest score across all rooms.
2. $1,500 to achieve the second highest final score across all rooms.
3. $500 to achieve the second highest final score across all rooms.

So you can win up to $4,000.

Note, you are not allowed to discuss any of the details of your competition room with other members of the community during the 2 weeks of the competition.

For any question please contact [PERSON CONTACT].

# Competition Rules for the Tournament w/ reserve

Dear [HANDLE],

Thanks for your interest in this competition. You have been randomly assigned to a room.

**In this room, the goal of the competition is to achieve the best final score, but only submissions with a final score of at least [SCORE THRESHOLD] can be awarded a prize.**

The top 2 coders to achieve the highest final scores will be awarded room prizes of $1,000 and $400 to the first and the second respectively. Only submissions that score at least [SCORE THRESHOLD] are eligible for prizes. No prizes will be awarded if no submissions reaches a score of at least [SCORE THRESHOLD].

You will be also competing for additional grand prizes across all rooms with the same competition style. Such grand prizes are distributed as follows:

1. $3,000 to the highest score of at least [SCORE THRESHOLD].
2. $1,500 to the second highest final score of at least [SCORE THRESHOLD].
3. $500 to the third highest score of at least [SCORE THRESHOLD].

So you can win up to $4,000.

Note, you are not allowed to discuss any of the details of your competition room with other members of the community during the 2 weeks of the competition.

For any question please contact [PERSON CONTACT].

Registration Survey

Thank you for your interest in this challenge. As part of the registration process we will be collecting participant information through this survey. The survey should take approximately 15 minutes to complete.

# Demographics

Age

Gender

Country of Origin

Country of Residence (the country in which you are planning to stay for the most part of the duration of the challenge)

Highest Academic Degree

Working or Student

# Risk aversion measure (List et al., 2014)

In this part of the survey, you will be making 10 choices between two options: OPTION A and OPTION B.

Imagine that each option is an hypothetical lottery, and every combination of the lotteries in Option A and B, are called Choice Pairs.

On the following Table, the first column denotes the Choice Pairs, numbered from 1 to 10. The second column presents the details of the lottery of Option A, and the third presents the details of the lottery of Option B.

Please have a look at the hypothetical lottery of Option A in Choice Pair 1.

Imagine the computer randomly selects a number from the range 1,2,3,..,10.

If the random number drawn is equal to “1”, this lottery pays $20; if the random number drawn is “2”, “3”, “4”, “5”, “6”, “7”, “8”, “9” or “10”, the lottery pays $16.

Similarly, the lottery in Option B of Choice Pair 1 pays $38.5 if the randomly drawn number is equal to “1”, and it pays $1.0 if the number drawn is “2”, “3”, “4”, “5”, “6”, “7”, “8”, “9” or “10”.

In the fourth column you can indicate which of the two lotteries in Choice Pair 1 you would prefer to participate in; the lottery as specified in Option A, or the lottery in Option B. The amounts are hypothetical, you are not entitled to earn any money after your responses.

After you have indicated whether you prefer to participate in the lottery of Option A or in that of Option B in Choice Pair 1, move to the second Choice Pair, and indicate whether you prefer Option A or B in that second Choice Pair.

In Choice Pair 2, the lottery in Option A pays €2.00 if the random number drawn is either “1” or “2”, and it pays €1.60 in case the random number drawn is equal to 3, 4, 5, ..., or 10. Similarly, the lottery in Option B of Choice Pair 2 pays €3.85 if the random number drawn is either “1” or “2”, and it pays €0.10 in case the random number drawn is equal to 3, 4, 5, ..., or 10.

Again, you can indicate in the fourth column which of the two lotteries in Choice Pair 2 you would prefer to participate in.

Note that the further down the screen you go, the larger the chances are of receiving the higher payoff in each of the two Options (€2.00 in Option A, and €3.85 in Option B), increases. In fact, in Choice Pair 10 you can receive €2.00 for certain if you choose Option A in that Choice Pair, or receive €3.85 with certainty if you choose Option B.

|  |  |  |  |
| --- | --- | --- | --- |
| Choice pair | Option A | Option B | Your Choice |
| 1 | Win $20 if rand =1  OR  win $16 if rand =3-10 | Win $38.5 if rand=1  OR  Win $1 if rand =2-10 | A or B |
| 2 | Win $20 if rand=1-2  OR  win $16 if rand =3-10 | Win $38.5 if rand=1-2  OR  Win $1 if rand =3-10 | A or B |
| 3 | Win $20 if rand=1-3  OR  win $16 if rand =4-10 | Win $38.5 if rand=1-3  OR  Win $1 if rand =4-10 | A or B |
| 4 | **…** | **…** | **…** |

# Time Availability

How many hours do you expect to be able to work on the solution of the problem in the next days? (look ahead a week, forecast how much you will be able)

* The first day of the competition [from 0 to 24]
* The second day of the competition [from 0 to 24]
* The third day of the competition [from 0 to 24]
* The remaining days of the competition

# In a typical day when you compete on a Top Coder MM, how many hours do you usually spend on the following activities?

Working on the solutions for the MM [from 0 to 24]

Education [from 0 to 24]

Job/Work [from 0 to 24]

Leisure [from 0 to 24]

Family [from 0 to 24]

Sleep [from 0 to 24]

# Risk aversion measure 2 (Dohmen et al., JEEA 2011)

How willing are you to take risks, in general?

[scale from 0 to 10]

Final Survey

Thank you for your interest in this study. As part of the research associated with the challenge we will be collecting participant information through this survey. The survey should take approximately 15 minutes to complete.

# What is your best estimate of the hours worked on the problem?

* Day 1(exact date) [scale from 0 to 24]
* Day 2 (exact date) …
* Day 3…
* […]
* Day 10

# How hard was to achieve a score of at least S in this competition?

[from very easy 0 to very hard 10]

# If the prize award was 2x how would your number of hours spent solving the problem have changed?

[scale in %]

# If the prize award was 0.5x how would your number of hours spent solving the problem have changed?

[scale in %]

# Give us your thoughts on competing in a race as opposed to a regular marathon match. Consider elements [engagement, planning required ahead of time, perception of competition, amount of effort exerted, quality of submissions, fairness]

[open answer]

# Please select why you dropped out from the competition (for those with no submissions):

* 1. I did not have the appropriate knowledge or skills to solve the problem.
  2. There were lots of strong competitors in my room.
  3. In the [race/tournament-min-req.,] the target score S was too high for me.
  4. I did not have time to participate in the competition due to other obligations.
  5. [I did not want to participate in a [race/tournament/tournament-min-req.]]
  6. I was not interested in topic of the competition.
  7. The awarded cash prize was too small for the effort required in solving the problem.
  8. Other [open]

# Could you list one ore more machine learning approaches you have used to solve the problem?

[open answer]