

Title: Yelp Dataset Challenge

Introduction:

Yelp connects people to great local businesses. To help people find great local businesses, Yelp engineers have developed an excellent search engine to sift through over 61 million reviews and help people find the most relevant businesses for their everyday needs.

Yelp is proud to introduce a deep dataset for research-minded academics from our wealth of data. If you've been looking for a rich set of data to train your models on and use in publications, this is it. Tired of using the same standard datasets? Want some real-world relevance in your research project? This data is for you!

How well can you guess a review's rating from its text alone? Can you take all of the reviews of a business and predict when it will be the most busy, or when the business is open? Can you predict if a business is good for kids? Has Wi-Fi? Has Parking? What makes a review useful, funny, or cool? Can you figure out which business a user is likely to review next? How much of a business's success is really just location, location, location? What businesses deserve their own subcategory (i.e., Szechuan or Hunan versus just "Chinese restaurants"), and can you learn this from the review text? What makes a tip useful? What are the differences between the cities in the dataset? There is a myriad of deep, machine learning questions to tackle with this rich dataset.

If you are a student and come up with an appealing project, you'll have the opportunity to win one of up to ten Yelp Dataset Challenge awards for \$5,000. Yes, that's \$5,000 for showing us how you use our data in insightful, unique, and compelling ways.

Additionally, if you win the challenge and then publish a research paper about your winning research in an peer-reviewed academic journal, then you'll be awarded an additional \$1,000 as recognition of your publication. If you are published, Yelp will also contribute up to \$500 to travel expenses to present your research using our data at an academic or industry conference.

There is no cost to enter, no purchase required, and you need not be a Yelp user or hold a registered Yelp account to participate.

Challenge Summary

- The Yelp Dataset Challenge begins August 1, 2014 and continues through December 31, 2014, however, Yelp may end the contest at any time at its sole discretion.
- Contest is open only to undergraduate and graduate university students in countries not restricted by the Office of Foreign Assets Control of the United States Department of the Treasury. Participants must be 18 or older and currently enrolled in an accredited degree program in any field.
- Access to the Yelp Dataset is governed by the separate Dataset Challenge Academic Dataset Agreement.
- To qualify for a \$5,000 Grand Prize, you must submit your code and a writeup of your project to Yelp via http://www.yelp.com/dataset_challenge before December 31, 2014. Up to ten individuals or team entrants may receive this Grand Prize. (See below for restrictions on team size and entry).
- Each Grand Prize winner may qualify for an additional \$1,000 Publication Prize if their work is

published by a peer-reviewed academic journal, and if published, an additional \$500 prize may be awarded to pay for attendance at an academic or industry conference to present your work.

- All prize awards are subject to Yelp's verification of your compliance with all stated rules.

Contest Rules, Terms and Conditions

Term

The Yelp Dataset Challenge ("Contest") entry period commences on August 1, 2014(00:00:00 UTC) and continues until December 31, 2014 (23:59:59 UTC) (the "Contest Term"). No additional registrations or Contest submissions will be accepted after the close of the Contest Term. Yelp reserves the right to cancel the Contest at any time, in its sole discretion.

Who can participate?

Contest participants may be individuals or teams. There is no restriction on team size. Each Participant consents to be bound by these Rules. For team participants, one person from the team must be designated as the team leader. A team leader will be the sole contact for all communications with Yelp regarding the contest. Individual team members, however, must each satisfy the contest requirement for individual participants. All Contest participants, whether entering the Contest as individuals or team members, must also each agree to the separate Dataset Challenge Academic Dataset Agreement.

To participate in the Contest, Participants must be at least eighteen (18) years old, and be enrolled as an Undergraduate or Graduate at an accredited university in any field.

Current and former employees, officers, interns, contractors and agents of Yelp, its subsidiaries, together with members of their immediate families (parent, child, sibling and spouse of each) and those living in their same households, or Yelp and any independent contractors supporting the Contest, are ineligible to enter and participate in the Contest or be awarded or retain any Contest prize.

Contest Participants may not be a national or permanent resident of any of the following countries: Burma, Cuba, Iran, North Korea, Sudan, or Syria. The Contest is void where prohibited or restricted by law. Yelp reserves the right to further limit, or restrict, participation in the Contest to any person at any time for any reason.

It is each Contest Participant's sole responsibility to review, understand and abide by their employer's or academic institution's policies, if any, regarding eligibility to participate in the Contest. If Yelp learns that a participant is in violation of their school's or employer's policies, they will be disqualified from the Contest and shall not be awarded, or allowed to retain, any Contest prizes. Yelp disclaims any and all liability or responsibility for disputes arising between a student or employee and their school or employer related to this Contest.

While individual team members may be members of multiple teams; teams with an identical set of members are not permitted.

Yelp may change any publicly displayed team data, including team names, if Yelp in its sole discretion deems it to be inappropriate.

Participants agree that Yelp may contact them regarding potential employment opportunities at Yelp.

All submissions to, and communications with, Yelp regarding the Contest must be in English. All entry information and submissions shall be deemed collected, stored, and judged in the United States.

Receipt and Judging of Qualifying Submissions

Submissions should reflect creative uses of the Dataset Challenge Academic Dataset. Submissions must be originally developed or implemented, and must not violate or infringe on any applicable law or regulation or third-party rights).

Participants may submit their project by uploading or submitting a link to a research paper or other document, blog, slide deck or video, or any other medium that sufficiently communicates their use of the Dataset to Yelp.

The submission, if written, must be written in English, with mathematical formulae as necessary. It must be written at a level sufficient for a practitioner in computer science to reproduce the results obtained by the participant, if any. Citations of previous related work should be referenced as appropriate.

Submissions must include a writeup of your project. Participants shall provide submissions to Yelp through the following website: yelp.com/dataset_challenge before December 31, 2014. Participants warrants that any source code is either fully or substantially developed and functions or will function as represented by the writeup or other description provided.

In order for Yelp to verify the submission, each Participant will grant to Yelp (including its affiliates and subsidiaries, employees, agents, and contractors), and any testing partners, an irrevocable, royalty free, fully paid up, worldwide non-exclusive license under the Participants' copyrights, patents or other intellectual property rights to use, review, assess, test and otherwise analyze the submissions and all their content in connection with this Contest.

All participants shall grant to Yelp (including its affiliates and subsidiaries, employees, agents, and contractors), an irrevocable, royalty free, fully paid up, worldwide non-exclusive license under the Participants' copyrights, patents or other intellectual property rights in their submission to reproduce, distribute, display, and create derivative works from the submission and also to make, have made, use, sell, offer for sale, and import products.

Award of Contest Prizes

Contest Prizes:

1. Grand Prizes: \$5,000 (USD)
2. Publication Prizes: \$1,000 (USD) (for Grand Prize winners only)
3. Conference Prizes: Up to \$500 (USD) for travel to, and attendance at, up to one (1) academic or industry conference to present your winning and published project (for Grand Prize and Publication Prize winners only)

A Participant, whether an individual or a team, is eligible to win and be awarded only one (1) Grand Prize of \$5,000 per team. Teams with multiple members are responsible for splitting the prize amongst

themselves. Up to 10 Participants may be chosen as Grand Prize winners, and Yelp reserves the right to choose fewer than 10 winners, including none, at its own discretion.

Only Grand Prize winners are eligible to receive Publication or Conference Prizes, and each Grand Prize winner is only eligible to receive up to one Publication Prize and up to one Conference Prize per Participant (in the case of teams, only one per team), even if the Grand Prize winner is published in multiple publications or attends multiple conferences.

For teams, any awarded prizes will be delivered to the team leader. We recommend teams determine ahead of time how prizes, if any, would be shared among team members. Participants, including teams, may only represent themselves, and may not represent any other entity (including any employer or academic institution).

In the event a potential winner cannot be reasonably notified, verified or confirmed based on two (2) attempts, a Contest Prize is declined, or a potential winner is not in compliance with these Rules or eligible for award of a Contest Prize for whatever reason, Yelp may allow the Contest to remain open and select and verify an alternate potential prize winner in accordance with these Rules, or to not select any winners, at its sole discretion.

Prizes will be distributed according to the following schedules:

(1) Grand Prizes will be distributed to winning Participants (to team leaders in the case of a team) in lump sum payments 4-6 weeks after Yelp has declared a winner.

(2) Publication Prizes will be distributed to winning Participants (to team leaders in the case of a team) in lump sum payments 4-6 weeks after Yelp has received sufficient proof of publication of research in an acceptable peer-reviewed academic journal. Publication in multiple journals does not warrant multiple prizes -- each winner is limited to one Publication Prize.

(3) Conference Prizes will be distributed to winning Participants (to team leaders in the case of a team) in lump sum payments 4-6 weeks after Yelp has received sufficient proof of attendance at a qualifying academic or industry conference, including evidence of payment of registration, travel, or lodging expenses incurred in connection with the conference.

Any costs or expenses associated with the acceptance and use of any Contest Prize monies are the sole responsibility of the Contest Prize winners. All Contest Prize details are at Yelp's sole discretion. A Contest Prize winner may make no substitutions or assignment of a Contest Prize other than as specified in these Rules.

Prize monies should be retained by individuals only in conformity with any applicable policies of their employers or academic institutions, regarding participation in and receipt of promotional consideration relating to the Contest and receipt and retention of Contest Prizes. If an employer's or school's policies are applicable, it is the Participant's sole and ultimate responsibility, in consultation with their employer or school, to determine how and if any Contest Prize will be retained and/or distributed and accounted for and Yelp assume no responsibility for the decisions made by such employers or schools regarding this issue.

General Contest Terms and Conditions

A. Delivery. A Contest participant is solely responsible for the equipment and Internet access required to participate in the Contest. Yelp will not be responsible for lost, late, incomplete, stolen, misdirected,

undelivered, delayed, inaccurate, garbled or illegible materials, entries, email or mail; or for any computer, telephone, cable, network, satellite, electronic or Internet hardware or software malfunctions, failures, connections, or availability, or for garbled or jumbled transmissions, or for service provider/Internet/Web site/use net accessibility or availability, traffic congestion, or unauthorized human intervention, or for inaccurate capture of any information, or the failure to capture any such information. Yelp is not responsible for any incorrect or inaccurate information, whether caused by printing errors, Site users, tampering, hacking, or by any of the equipment or programming associated with or utilized in the Contest, and is not responsible for any error, omission, interruption, deletion, defect, delay in operation or transmission, communications line failure, theft or destruction or unauthorized access to the Contest Site or any communications means or for any other error or failure of any kind in connection with or relating to this Contest, whether caused by any human, typographical, printing, mechanical, computer or other means, including without limitation errors or failures which may occur in connection with the processing or storing of information, the administration of the Contest, the uploading, processing validating, posting and/or judging of entries, the announcement of the Contest Prizes or in any Contest-related materials. Yelp is not responsible for injury or damage to any participants' or to any other person's computer related to or resulting from participating in this Contest or downloading materials from or use of the Site. Yelp shall have no obligation to accept or consider entries delayed past the submission deadline.

B. Entries; Verification/Audit; Participant Cooperation. All activities relating to Participant's participation in the Contest and material submitted are subject to verification and/or auditing for compliance with these Rules and Participants agree to reasonably cooperate with Yelp concerning verification and/or auditing. In the event that Contest verification activity or an audit evidences non-compliance with the Rules or official Contest communications, as determined in Yelp's reasonable discretion, a Participant's continuing participation in any aspect of the Contest may be suspended or terminated at any time during or after the Contest.

Yelp reserves the right, in its sole discretion, to suspend or cancel the Contest. In the event the Contest is cancelled, Yelp may, at its sole discretion, determine winners, if any, from among all eligible entries received up to the time of such cancellation, but it has no obligation to do so.

C. Participants' Affidavits, Contest Consent and Releases. As a prior condition of winning the Contest and receipt and acceptance of a Contest Prize, a participant may be required to complete, sign and return a non-exclusive grant of license to Yelp, an Affidavit of Eligibility and Liability Release and where lawful, a Publicity and Submission Release and License ("Publicity Release"). The Publicity Release will include consent by a winning participant to use their name, picture, voice, likeness, and the name, description and Application comprising winner's entry, worldwide in perpetuity, for the purposes of Contest-related advertising, promotion and publicity, and additionally including those specific purposes disclosed on the Site, without further compensation, notification or permission, unless prohibited by law. Failure to complete and sign these documents within two (2) weeks may disqualify that entry and additional qualifying entries will be considered.

In addition, winning participants may be required to complete applicable U.S. tax withholding-related documentation and, if necessary, provide proof of foreign status, as beneficiaries of U.S.-sourced promotional consideration.

D. Tax Consequences. Participation and receipt of benefits from the Contest will have tax consequences. Participants should speak with their tax advisors prior to participation and receipt and acceptance of Contest Prizes to prevent any undesired results. Should there be any local tax liability for participation in

the Contest or the receipt of a Contest Prize, or costs or expenses relating to participation in the Contest or for the use and enjoyment of any Contest Prizes, or for any other reason, such taxes and/or expenses are the sole responsibility of the participant. Yelp will issue Contest-related tax reporting, as may be required by applicable laws and regulations. As a condition of winning the Contest and receipt and use of any Contest Prize, participants agree to self-report to applicable local taxing authorities, as may be required by local laws.

E. Rules Amendment and Interpretation. The Contest and its continuing terms and conditions, benefits and participation are offered to participants at the sole discretion of Yelp. Yelp reserves the right to amend or interpret these Rules or any other official Contest communications and any element or elements arising under or relating to the Contest at any time, upon published notice to participants on the following site: yelp.com/dataset_challenge. A participant shall be deemed to have notice of any such amendments or interpretations upon publication of the same and shall be deemed to have accepted such amendments or interpretations by virtue of a participant's continuing participation in the Contest. Should a participant not wish to continue to participate in the Contest pursuant to the prevailing Rules, as amended or interpreted, a participant may terminate participation in the Contest after providing a submission by notifying Yelp in writing.

F. Rule Compliance Interpretation; Adjudication and Remedial Process. Yelp reserves the right to make all decisions relating to the Contest at its sole discretion, and such decisions by Yelp are final and binding. Participants in the Contest agree that: (1) any and all disputes, claims, and/or causes of action arising out of or connected with this Contest, or any benefits received, or the administration of the Contest, shall be resolved individually, without resort to any form of class action, and exclusively by arbitration, to take place in the City and County of San Francisco in the State of California pursuant to the Rules of the American Arbitration Association, then effective, (2) any and all claims, judgments and awards shall be limited to actual out-of-pocket costs incurred, including costs associated with participating in the Contest and in no event shall Participant be entitled to receive attorneys' fees or other legal costs; (3) under no circumstances will a Participant be permitted to obtain awards for and Participants hereby waive all rights to claim punitive, incidental and consequential damages and any other damages, other than for actual out-of-pocket expenses, and; (4) any and all disputes, claims, and causes of action arising out of or relating to the Competition, the Data or any prize awarded must be brought within one (1) year after the occurrence of the event giving rise to such claim (if such claim is not filed, then that claim is permanently barred). Except where prohibited, all issues and questions concerning the construction, validity, interpretation and enforceability of these Rules, or the rights and obligations of Participant(s) and Yelp in connection with the Contest, shall be governed by, and construed in accordance with, the laws of the United States of America and the State of California, without regard for conflicts of law doctrine of any jurisdiction, and subject to resolution only in the City and County of San Francisco, State of California, United States of America. By entry and participation in the Contest, Participants consent to personal jurisdiction in the City and County of San Francisco, State of California, United States of America, and agree that the City and County of San Francisco, State of California, United States of America will be the exclusive forum and venue for any dispute resolution. Yelp's failure to exercise or enforce any right or provision under these Rules will not constitute a waiver of such right or provision and will not be deemed a waiver of any further rights hereunder.

G. NO WARRANTY. YELP, ITS AFFILIATED COMPANIES AND ALL OF THEIR RESPECTIVE OFFICERS, DIRECTORS, EMPLOYEES, REPRESENTATIVES AND AGENTS (COLLECTIVELY, THE "RELEASED PARTIES") EXPRESSLY DISCLAIM ANY AND ALL WARRANTIES OR CONDITIONS OF ANY KIND (WHETHER EXPRESS, IMPLIED, STATUTORY OR OTHERWISE), INCLUDING BUT NOT

LIMITED TO, IMPLIED WARRANTIES OR CONDITIONS OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT. UNDER NO CIRCUMSTANCES SHALL RELEASED PARTIES BE HELD RESPONSIBLE OR LIABLE FOR A PARTICIPANT'S USE OF THE INFORMATION AND/OR PRODUCTS PROVIDED AND/OR MADE AVAILABLE THROUGH THE CONTEST OR FOR ERRORS OR ANOMALIES RESULTING IN THE UNINTENDED OR ERRONEOUS PARTICIPATION, AWARD OF CONTEST PRIZES OR OTHER BENEFITS UNDER THE CONTEST TO PARTICIPANTS. RELEASED PARTIES OFFER NO ASSURANCES, GUARANTEES OR WARRANTIES OR CONDITIONS THAT THE CONTEST OR CONTEST WEBSITES WILL BE UNINTERRUPTED OR ERROR-FREE AND DOES NOT GUARANTEE THE ACCURACY OR RELIABILITY OF ANY INFORMATION OBTAINED THROUGH THE CONTEST. RELEASED PARTIES ASSUME NO RESPONSIBILITY FOR ANY COMPUTER-RELATED DAMAGES DUE TO DOWNLOADING MATERIALS. RELEASED PARTIES WILL NOT BE LIABLE, AND ARE NOT RESPONSIBLE, FOR DAMAGES OF ANY KIND RELATED TO A PARTICIPANT'S PARTICIPATION OR INABILITY TO PARTICIPATE IN THE CONTEST, WHETHER THE DAMAGES ARE DIRECT, INDIRECT, INCIDENTAL, SPECIAL OR CONSEQUENTIAL. FURTHER, BY PARTICIPATING IN THE CONTEST AND/OR ACCEPTING A CONTEST PRIZE, A PARTICIPANT AGREES THAT RELEASED PARTIES SHALL NOT BE LIABLE FOR, AND WILL BE HELD HARMLESS BY PARTICIPANT AGAINST, ANY LIABILITY FOR ANY DAMAGE, INJURY OR LOSS TO PERSON (INCLUDING DEATH) OR PROPERTY DUE IN WHOLE OR IN PART, DIRECTLY OR INDIRECTLY, TO ACCEPTANCE, POSSESSION, USE OR MISUSE OF A CONTEST PRIZE, PARTICIPATION IN ANY CONTEST PRIZE-RELATED ACTIVITY, RELEASED PARTIES' USE OF ANY TENDERED ENTRY OR PARTICIPATION IN THIS CONTEST. SOME STATES DO NOT ALLOW THE LIMITATION OR EXCLUSION OF LIABILITY FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES; THEREFORE THE ABOVE LIMITATIONS OR EXCLUSIONS MAY NOT APPLY.

H. Indemnification. Participants agree to release, defend, indemnify and hold harmless Released Parties from and against, and accept all responsibility of any kind, including but not limited to financial, for any liability, claims, losses, damages or proceedings, including but not limited for death, (including reasonable attorneys' fees) relating to any actions taken by a participant, including any submissions, or anyone else using a participant's e-mail or team password or otherwise purporting to act on participant's behalf in regard to the Contest (whether or not such use occurred with or without your permission). Participant also agrees to release, defend, indemnify and hold harmless the Released Parties from any and all liability, claims, losses, damages or proceedings, including but not limited for death (including reasonable attorneys' fees) relating to Prizes and any other matter in connection with a participant's participation in the Contest or Yelp's or its designees or Yelp's or its designees use of his/her Submission.

These rules, terms and conditions specified here comprise the Contest's Official Rules and, together with the Dataset Challenge Academic Dataset Agreement and official Contest communications or publications by Yelp, shall govern and apply to all participation and activity arising out of and relating to the Contest and use of the Site. These Rules cannot be modified or superseded except by Yelp, at its sole discretion.