

NATIONAL DATA SCIENCE CHALLENGE 2019

COMPETITION GUIDE - BEGINNER CATEGORY

We will be using Kaggle for the NDSC. Kaggle is a platform for predictive modelling and analytics competitions in which statisticians and data miners compete to produce the best models for predicting and describing the datasets uploaded by companies and users.

Final submission deadline of the project - 23 March 2019, 11.59pm (SGT)

Note that the competition organisers reserve the right to update the competition timeline if they deem it necessary.

To get started, please register on Kaggle.

- 1) Create an account on [Kaggle](#)
- 2) Activate your Kaggle account.

After registering for an account, please **click [here](#)** to join the National Data Science Challenge - Beginner Category on Kaggle.

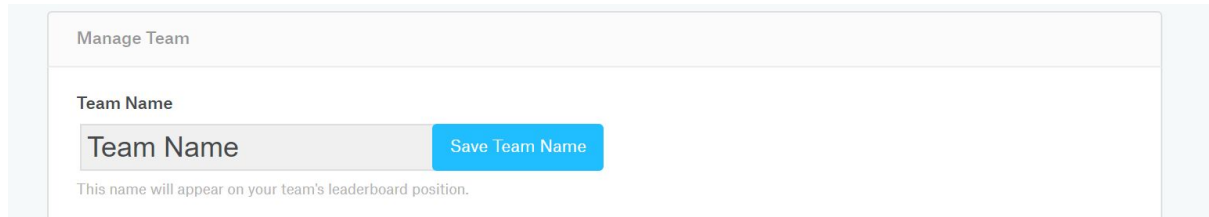
Further Information

The following Annexes are enclosed for your information:

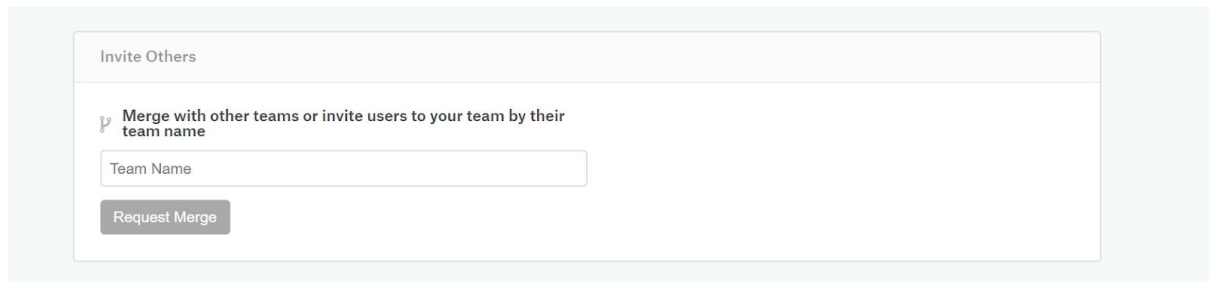
- Annex A - Team Formation
- Annex B - Deliverables
- Annex C - Submission on Kaggle
- Annex D - Terms and Conditions

TEAM FORMATION

- 1) Read the competition rules carefully.
- 2) If you agree to be bound to the competition rules, click "I Understand and Accept".
- 3) The team leader can go to the "Team" tab (In the main page of the competition, the tabs are: Overview, Data, Kernels, Discussion, Leaderboard, Rules, Team)



- 4) Under "Merge with other teams", enter the name of your team members. Please ensure that all the team members have registered for a Kaggle account.



- 5) Form your team.
- 6) **Please ensure that the team name on Kaggle is the same as the team name you registered for via the NDSC website.** Failure to comply will result in disqualification of the challenge.

For more information on Team Formation, please click [here](#).

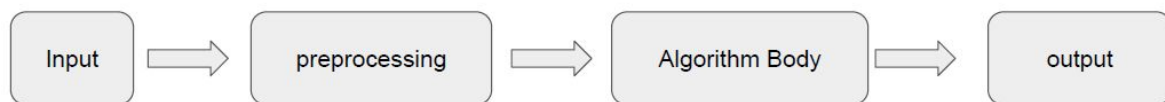
DELIVERABLES**1) Experimental Results**

Please upload in .csv file on Kaggle.

2) Supplementary Material

The supplementary material should be written in English and submitted in .doc or .pdf file (maximum 4 pages) via email.

The supplementary material should include the algorithm pipeline, result analysis and summary based on your methodology.

Algorithm Pipeline

- You may highlight the innovative idea about this project.
- You may provide formulation/ theoretical proof to support your idea.
- You may give analysis on your experiments.
- You may want to draw conclusion based on your experiments results and analysis.
- You can share your reflections.

3) Source Code

The source code should be wrapped in .zip file.

Please submit the **Experimental Results** on Kaggle.

Please submit **both the Supplementary Material and Source Code** to ndsc@shopee.com in the following format:

- Subject title: [Beginner] Final Submission - (Team Name)
- Copy all team members in the email
- Deadline: **23 March 2019, 11:59 pm SGT**

Note: Winners are required to provide a detailed report of their solution. Reports need to be written in English. Reports should be approximately four (4) pages long, and include technical details acceptable for an academic workshop. Teams must provide their report to Shopee no later than two (2) weeks after the Submission deadline.

SUBMISSION ON KAGGLE**For scripts**

Step 1: Save your submissions file.

```

✓ Run Successful - Run Time: 0:05 View Completed Script (Exit Editor) »
1 # This Python 3 environment comes with many helpful analytics libraries installed
2 # It is defined by the kaggle/python docker image: https://github.com/kaggle/docker-python
3 # For example, here's several helpful packages to load in
4
5 import numpy as np # linear algebra
6 import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
7
8 # Input data files are available in the "../input/" directory.
9 # For example, running this (by clicking run or pressing Shift+Enter) will list the files in the input directory
10
11 from subprocess import check_output
12 print(check_output(["ls", "../input"]).decode("utf8"))
13
14 # Any results you write to the current directory are saved as output.
15
16 sub = pd.read_csv('../input/gender_submission.csv')
17
18 sub.to_csv('sample_submission.csv', index=False)

```

Step 2. Click “Submit to Competition”.

Data **Output** Log Info Version 3

sample_submission.csv

sample_submission.csv

2.77 KB

Download

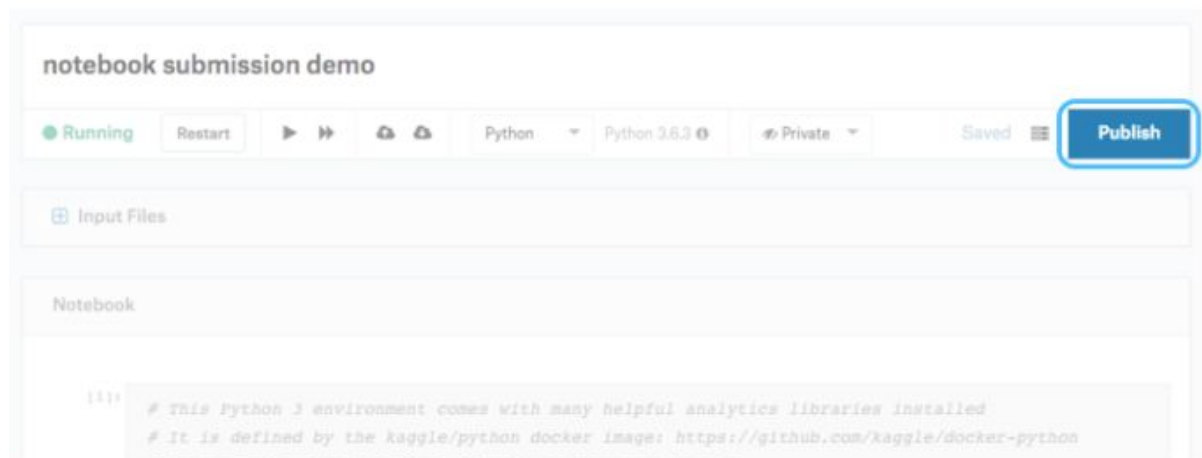
Submit to Competition

Preview (first 100 rows)

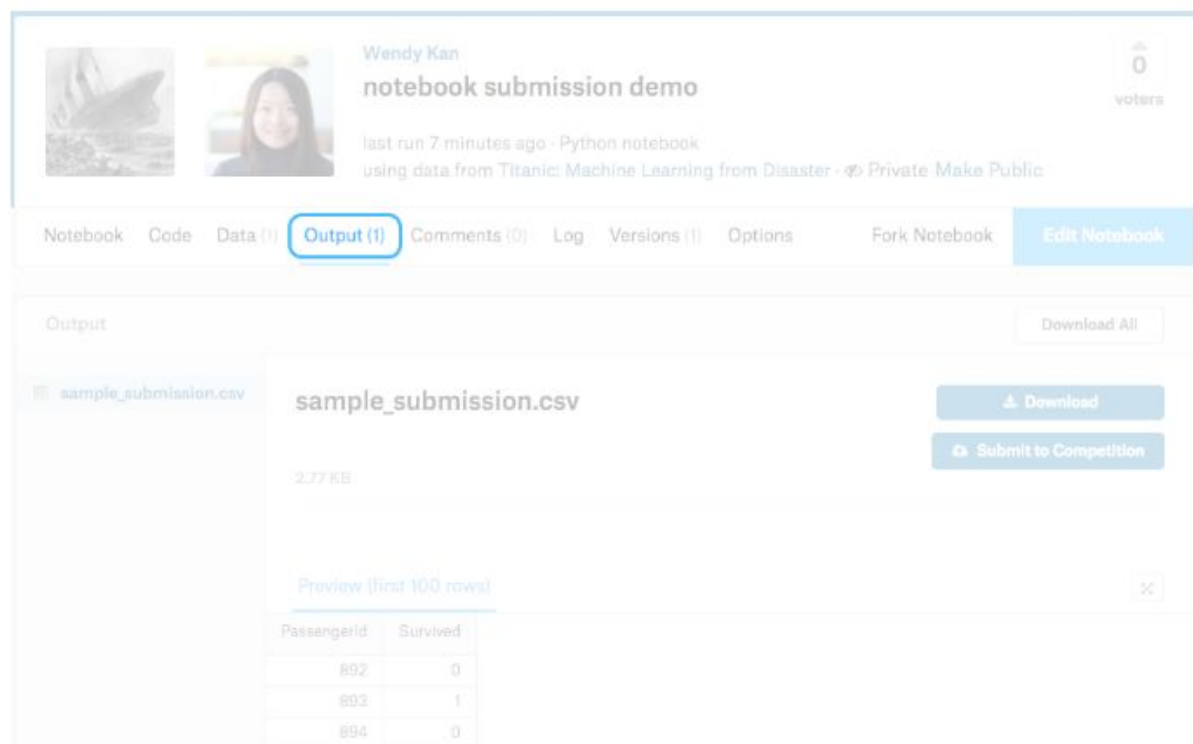
PassengerId	Survived
892	0
893	1
894	0
895	0
896	1
897	0

For notebooks:



Step 1: “Publish” (this will not make your kernel public).



Step 2: Click the “Output” tab.



Step 3: Click “Submit to Competition”.



Wendy Kan

notebook submission demo

last run 7 minutes ago · Python notebook
using data from Titanic: Machine Learning from Disaster · Private Make Public

0 voters

Notebook

Code

Data (1)

Output (1)

Comments (0)

Log

Versions (1)

Options

Fork Notebook

Edit Notebook

Output

Download All

sample_submission.csv

sample_submission.csv

2.77 KB

Download

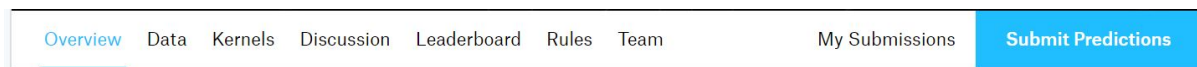
Submit to Competition

Preview (first 100 rows)

PassengerId	Survived
892	0
893	1
894	0

For CSV file

1. Click 'Submit Predictions'



2. Upload submission file. Submission should be in CSV format.

Step 1

Upload submission file



File Format

Your submission should be in CSV format. You can upload this in a zip/gz/rar/7z archive, if you prefer.

Number of Predictions

We expect the solution file to have 123623 prediction rows. This file should have a header row. Please see sample submission file on the [data page](#).

3. Describe submission.

Step 2

Describe submission

B **I** | 🔗 “ </> 🖼️ | ☰ ☷ H 📏 | ↺ ↻ 📄 Styling with Markdown supported

Briefly describe your submission.

4. Click 'Make Submission'.

For more information on Submission, please click [here](#).

TERMS AND CONDITIONS

1. Introduction

Shopee Singapore Private Limited ("Shopee") is organising a competition "National Data Science Challenge 2019" (the "Competition") to be held from 23 February 2019 until 23 March 2019 ("Competition Period").

The objective of the Competition is to provide an environment for the development of creative new ideas in Data Science.

2. Competition Details

The Competition consists of two phases:

2.1. Phase 1

Competition participants ("Participants") will submit program algorithms and/or codes (each submission, an "Entry") for judging through Kaggle at <http://www.kaggle.com/c/ndsc-beginner>

2.2. Phase 2

Each Entry will be scored and ranked by the evaluation metric set forth on Kaggle at <https://www.kaggle.com/c/ndsc-beginner#evaluation>. During the Competition Period, the current ranking for each Entry will be visible via the leader boards on Kaggle at <https://www.kaggle.com/c/ndsc-beginner/leaderboard>.

In the event of a tie, the Entry submitted first to the Competition will be the winner.

3. Prizes

First Prize: Singapore Dollars (SGD) Three Thousand Dollars (SGD \$3,000);
Second Prize: Singapore Dollars (SGD) Two Thousand Dollars (SGD \$2,000);
Third Prize: Singapore Dollars (SGD) One Thousand Dollars (SGD \$1,000); and
Most Creative Team Prize: Singapore Dollars (SGD) Eight Hundred Dollars (SGD \$800)

4. Participation Requirements

By submitting an Entry and participating in this Competition, the Participant agrees to comply with these Competition Terms and Conditions at all times.

4.1. One account per Participant

Each Participant shall be limited to only one Kaggle account for the purposes of submitting Entries to the Competition.

4.2. No sharing outside of teams

Privately sharing algorithms, codes and/or data for the purposes of the

Competition, outside of the Participant's team is not permitted.

4.3. No hand labelling of data

Participants are not allowed to hand label Shopee Data (defined herein) provided to the Participants during the Competition Period. Shopee reserves the right to test the teams' source codes in their respective Entries to verify the authenticity of such Entry. Shopee reserves the right to revoke the teams' ranking if hand labelling is detected in any Entries of such team.

4.4. Team Size Limits

The team size consists of a minimum of 3 to a maximum of 4 Participants per team. Individuals who registers for the Competition and does not belong to a team shall be allocated to a team at Shopee's sole discretion.

4.5. Eligibility

Employees of Sea Limited are not eligible to participate in the challenge in the interest of fairness.

4.6. Team Mergers

Mergers between teams are not allowed.

4.7. External Data

Usage of publicly, freely available external data in the teams' Entry is permitted.

5. Entry Requirements

The Competition is open only to registered Participants and Participants can register for the Competition via the Competition Website.

Shopee shall announce the rules and judging criteria of the Competition and any such rules and judging criteria shall be deemed to be incorporated into these Competition Terms and Conditions.

Teams may submit a maximum of two (2) Entries per day and teams can select two (2) final Entries from such teams' past submissions during the Competition Period for judging.

Shopee reserves the final right to reject any Entries in its sole discretion.

6. Obligations of the Participants

- 6.1.** Participant grants Shopee the right to use the Participant's name in connection with such Participant's submitted Entry as well as in connection with any marketing, advertising or promotional material related to the Competition without requiring any further consent or compensation from Shopee.

- 6.2.** Participant represents and warrants (where applicable) that it is duly organized, validly existing, and in good standing in the jurisdiction it is formed, that it will comply with all applicable laws in performing under these Competition Terms and Conditions and that it has full capacity, authority and legal right to agree to these Competition Terms and Conditions.
- 6.3.** Participants agree that they shall not:
- 6.3.1.** use or access datasets provided by Shopee for the purpose of this Competition (“Shopee Data”) for any commercial, gambling, or illegal purpose;
 - 6.3.2.** copy, modify, reproduce, license, sublicense, distribute, transmit, republish, exhibit, or display the Shopee Data (except to the extent copying, modifying, reproducing or displaying is necessary and directly related to the Participants’ participation in the Competition);
 - 6.3.3.** remove, obscure, or modify any copyright notices or other notices, designations, or attributions within the Shopee Data; or
 - 6.3.4.** use any Shopee Marks or any copyright or other proprietary rights associated with the Shopee Data for any purpose without the express written consent of Shopee “Shopee Marks” means all names, trademarks, logos, symbols, designations, emblems, designs, and trade dress of the Shopee, together with all identifications, labels, insignia or indicia thereof.
- 6.4.** The winning team of the Competition shall license the winning Entry and the source code used to generate the Entry under an Open Source Initiative-approved license that in no event limits commercial use of such code or model containing or depending on such code.

7. Intellectual Property Rights

- 7.1.** All intellectual property and proprietary rights in each Participant’s Entry shall belong to Participant, save for Shopee’s intellectual property embedded in such Entry which shall remain the sole property of Shopee.
- 7.2.** Participant grants to Shopee, its and its subsidiaries’ and affiliates’ agents, contractor and partners, an irrevocable, non-exclusive, royalty-free, worldwide, perpetual and sub-licensable license to reproduce, distribute, transmit, create derivative works of, publicly display and publicly perform the Entry.
- 7.3.** Shopee grants to Participants for the Competition Period a non-exclusive, non-transferrable and revocable license to use Shopee’s Data solely in connection with the Competition. Participant shall not, without the prior written consent of Shopee, alter or modify or create derivative works of the Shopee Data or any of Shopee’s intellectual property except as expressly set forth in these Competition Terms and Conditions.

7.4. Nothing in these Competition Terms and Conditions is intended to grant Participant any rights to use any of Shopee's intellectual property, except as specified in these Competition Terms and Conditions.

7.5. Participant warrants that (a) its Entry is their original work and (b) Participant has all rights and ownership or is a licensed user of all algorithms, code and/or data used within the Entry.

8. Indemnification and Limitation of Liability

Participant shall fully indemnify Shopee and hold Shopee harmless from and against any and all claims, actions, losses, damages, liabilities, costs, and expenses, including attorneys' fees and other legal expenses arising directly or indirectly from or in connection with (a) any breach by Participant of these Competition Terms and Conditions; (b) any failure of Participant to perform its obligations under these Competition Terms and Conditions in compliance with all applicable laws; (c) any violation of any rights of any third party related to Participant's Entry; and (d) Participant's fraud, negligence or wilful misconduct.

9. Termination

Shopee reserves the right to terminate the Competition at its sole discretion and at any time. The exercise by Shopee of its right to terminate the Competition shall not entitle Participant to any fees, damages, claims for expenses or lost profits, or any other recourse in law or in equity in respect of such termination.

10. Force Majeure

Shopee shall not be liable to Participant or be deemed to be in breach of these Competition Terms and Conditions by reason of any delay in performing or failure to perform any of its obligations, if such delay or failure was due to any event or circumstances beyond Shopee's reasonable control.

11. Decision

Shopee's decision on all matters regarding the Competition is final and binding on all Participants and Shopee shall not be required to provide any reason or explanation for its decision. Any reason or explanation provided by Shopee is for information only, shall not be binding on Shopee and does not constitute any representation, warranty or undertaking as to future action or otherwise.

12. Confidential Information

12.1. "Confidential Information" means all information that Shopee deems to be confidential,

12.2. Participant will: (a) treat as confidential all Confidential Information of Shopee; (b) not disclose that Confidential Information to any third party; (c) not use that Confidential Information except in connection with performing its obligations or exercising its rights under these Competition Terms and Conditions.

13. Personal Data

Participant agrees to Shopee's collection, use and disclosure of its personal data for the purposes of contacting Participant or its authorised representative by phone call, SMS or email in relation to the Competition.

14. General

14.1. Amendments

Shopee reserves the right to amend these Competition Terms and Conditions from time to time at its sole discretion.

14.2. No Waiver

A failure by Shopee to exercise or enforce any rights conferred upon it by these Competition Terms and Conditions shall not be deemed to be a waiver of any such rights or operate so as to bar the exercise or enforcement thereof at any subsequent time or times. Any express statement of a right of Shopee herein is without prejudice to any other right of Shopee expressly stated herein or arising at law.

14.3. Severability

If at any time any provision of these Competition Terms and Conditions shall be or shall become illegal, invalid or unenforceable in any respect, the legality, validity and enforceability of the remaining provisions of these Competition Terms and Conditions shall not be affected or impaired thereby, and shall continue in force as if such illegal, invalid or unenforceable provision was severed from these Competition Terms and Conditions.

14.4. Rights of Third Parties

A person or entity who is not a party to these Competition Terms and Conditions shall have no right to enforce or to enjoy the benefit of any term of these Competition Terms and Conditions.

14.5. Governing Law

These Competition Terms and Conditions shall be governed by and construed in accordance with the laws of Singapore, and each Participant submits to the non-exclusive jurisdiction of the courts of Singapore.