# Predict The News Category Hackathon

From the beginning, since the first printed newspaper, every news that makes into a page has had a specific section allotted to it. Although pretty much everything changed in newspapers from the ink to the type of paper used, this proper categorization of news was carried over by generations and even to the digital versions of the newspaper. Newspaper articles are not limited to a few topics or subjects, it covers a wide range of interests from politics to sports to movies and so on. For long, this process of sectioning was done manually by people but now technology can do it without much effort. In this hackathon, Data Science and Machine Learning enthusiasts like you will use Natural Language Processing to predict which genre or category a piece of news will fall in to from the story.

**Size of training set:**7,628 records  
**Size of test set:**2,748 records

**FEATURES:**

STORY:  A part of the main content of the article to be published as a piece of news.  
SECTION: The genre/category the STORY falls in.

There are four distinct sections where each story may fall in to. The Sections are labelled as follows :

Politics: 0  
Technology: 1  
Entertainment: 2  
Business: 3

Hackathon rules

**Eligibility:**+

Hackathons are open to all registered users at www.machinehack.com.a participant must be 18 years or older.+

**Participation:**+

The competition named above is a skills-based competition to promote and further the field of data science. You must register via the competition website to enter. Your competition submissions (“submissions”) must conform to the requirements set forth on the competition website. Your submissions will be scored based on the evaluation metric described on the competition website. Subject to compliance with the competition rules, prizes described on the competition website, if any, will be awarded to participants with the best scores, based on the merits of the data science models submitted. See below for the complete competition rules.+

**Determining winners:**+

Each submission will be scored and ranked by the evaluation metric set forth on the machine hack website. During the hackathon, the current ranking will be visible on the machine hack website’s leader board. The potential winner(s) are determined solely by the leader board ranking on the private leader board as set forth on the website, subject to compliance with these rules.+

In the event of a tie, the submission that was entered first to the hackathon will be the winner. In the event a potential winner is disqualified for any reason, the submission that received the next highest score rank will be chosen as the potential winner.+

Machine hack will notify the potential winner(s) by email. If a potential winner does not respond to the notification attempt within five (5) days from the first notification attempt, then such potential winner will be disqualified and an alternate potential winner will be selected from among all eligible entries received based on the judging criteria described herein.+

The winner’s list will be publicly displayed at machinehack.com. Determinations of hackathon sponsor are final and binding.+

**Prizes:**+

All prizes are subject to hackathon sponsor’s review and verification of the participants eligibility and compliance with these rules, and the compliance of the winning submissions with the submissions requirements. In the event that the submission demonstrates non-compliance with these competition rules, competition sponsor may at its discretion take disqualify the submission(s).+

Transfer or assignment of a prize is not allowed.+

If a team wins a monetary prize, the prize amount would be evenly shared between the team members.+

**Representation :**+

If you are entering as a representative of a company, educational institution or other legal entity, or on behalf of your employer, these rules are binding on you, individually, and/or the entity you represent or are an employee. If you are acting within the scope of your employment, as an employee, contractor, or agent of another party, you warrant that such party has full knowledge of your actions and has consented thereto, including your potential receipt of a prize. You further warrant that your actions do not violate your employer’s or entity’s policies and procedures.+

The hackathon sponsor reserves the right to verify eligibility and to adjudicate on any dispute at any time. If you provide any false information relating to the hackathon concerning your identity, residency, mailing address, telephone number, email address, ownership of right, or information required for entering the hackathon, you may be immediately disqualified from the hackathon.+

**Taxes.**+

All taxes imposed on prizes are the sole responsibility of the winners.+

**Governing Law:**+

The hackathon shall be governed by the laws of India. The courts of law at Bangalore Karnatka shall have exclusive jurisdiction over any disputes arising under this agreement. Hackathon sponsor reserves the right to disqualify any entrant from the hackathon if, in hackathon sponsor’s sole discretion, it reasonably believes that the entrant has attempted to undermine the legitimate operation of the hackathon by cheating, deception, or other unfair playing practices or abuses, threatens or harasses any other entrants, hackathon sponsor or Analytics India Magazine Pvt ltd.+

**Publicity.**+

By accepting a prize, you agree that hackathon sponsor, Analytics India magazine and its agencies may use of your name and/or likeness, for advertising and promotional purposes without additional compensation, unless prohibited by law.+

**Liability**+

Hackathon  Entities are not responsible for any malfunction of the Hackathon Website or any late, lost, damaged, misdirected, incomplete, illegible, undeliverable, or destroyed Submissions or entry materials due to system errors, failed, incomplete or garbled computer or other telecommunication transmission malfunctions, hardware or software failures of any kind, lost or unavailable network connections, typographical or system/human errors and failures, technical malfunction(s) of any telephone network or lines, cable connections, satellite transmissions, servers or providers, or computer equipment, traffic congestion on the Internet or at the Hackathon Website, or any combination thereof, which may limit an entrant’s ability to participate.+