Knowledge • 136 teams

Transfer Learning on Stack Exchange Tags

Fri 28 Oct 2016 Sat 25 Mar 2017 (3 months to go)

Dashboard

Home
- Data
- Make a submission

Information
- Description
- Evaluation
- Rules

Forum

Kernels
- New Script
- New Notebook

Leaderboard

My Submissions

Public Leaderboard

1. Arjun Variar
2. Thalles Domician [BCC]
3. Istvan Nagy
4. Karol Grzegorczyk
5. Aleksei Solovev
6. connectwithghosh
7. Yunfeng Zhu
8. binwang
9. LinWei
10. NTU_b02901004_hail jas

159 Kernels

Tag to Title/Content Matching in R - First Attempt
6 Votes / 14 days ago / R

tf-idf
5 Votes / 21 days ago / Python

Supervised learning / Binary relevance
2 Votes / 2 days ago / Python

Tags Exploration (distribution and shared tags)
2 Votes / 2 days ago / Python

Frequent Words Model
3 Votes / 23 days ago / Python

Word Clouds

Competition Details » Get the Data » Make a submission

Predict tags from models trained on unrelated topics

What does physics have in common with biology, cooking, cryptography, diy, robotics, and travel? If you answered "all pursuits are governed by the immutable laws of physics" we'll begrudgingly give you partial credit. If you answered "all were chosen randomly by a scheming Kaggle employee for a twisted transfer learning competition", congratulations, we accept your answer and mark the question as solved.

In this competition, we provide the titles, text, and tags of Stack Exchange questions from six different sites. We then ask for tag predictions on unseen physics questions. Solving this problem via a standard machine approach might involve training an algorithm on a corpus of related text. Here, you are challenged to train on material from outside the field. Can an algorithm learn appropriate physics tags from "extreme-tourism Antarctica"? Let's find out.

Kaggle is hosting this competition for the data science community to use for fun and education. This dataset originates from the Stack Exchange data dump.

Started: 5:44 pm, Friday 28 October 2016 UTC
Ends: 11:59 pm, Saturday 25 March 2017 UTC (148 total days)
Points: this competition does not award ranking points
Tiers: this competition does not count towards tiers