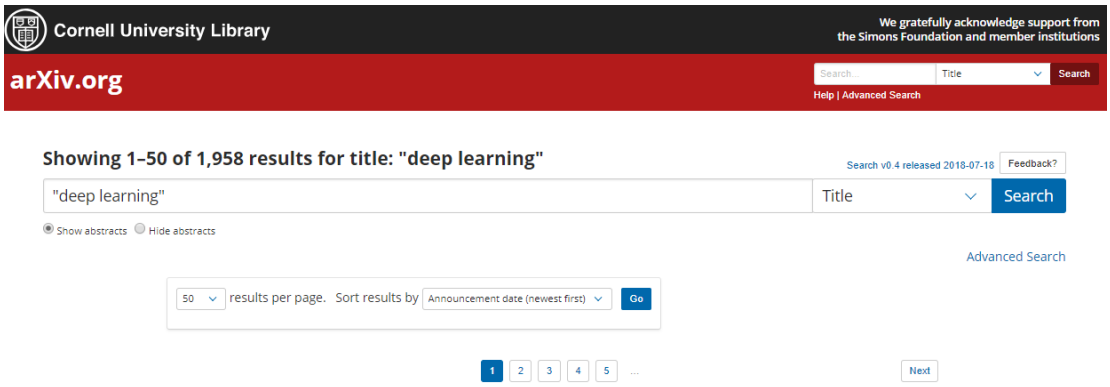
Assignment 4: Web scraping from arXiv.org

Step 1: visit the website https://arxiv.org

Step 2: provide "deep learning" in the search form with "Title" option



Step 3: Modify the provided script "04_3_Web Scraping_arXiv papers" as follows:

Step 3-1: Read the url of each paper from the first page to the last page

1. arXiv:1811.04017 [pdf, other] cs.LG
A generic framework for privacy preserving deep learning
Authors: Theo Ryffel, Andrew Trask, Morten Dahl, Bobby Wagner, Jason Mancuso, Daniel Rueckert, Jonathan Passerat-Palmbach
Abstract: We detail a new framework for privacy preserving deep learning and discuss its assets. The framework puts a premium on ownership and secure processing of data and introduces a valuable representation based on chains of commands and tensors. This abstraction allows one to implement complex privacy preserving constructs such as Federated Learning, Secure Multiparty Computation, and Differential Priv... ▽ More
Submitted 9 November, 2018; **originally announced** November 2018.
Comments: PPML 2018, 5 pages

2. arXiv:1811.03970 [pdf, other] cs.IR
Looking Deeper into Deep Learning Model: Attribution-based Explanations of TextCNN
Authors: Wenting Xiong, Iftitahu Ni'mah, Juan M. G. Huesca, Werner van Ipenburg, Jan Veldsink, Mykola Pechenizkiy
Abstract: Layer-wise Relevance Propagation (LRP) and saliency maps have been recently used to explain the predictions of Deep Learning models, specifically in the domain of text classification. Given different attribution-based explanations to highlight relevant words for a predicted class label, experiments based on word deleting perturbation is a common evaluation method. This word removal approach, howev... ▽ More
Submitted 8 November, 2018; **originally announced** November 2018.

3. arXiv:1811.03962 [pdf, other] cs.LG
A Convergence Theory for Deep Learning via Over-Parameterization
Authors: Zeyuan Allen-Zhu, Yuanzhi Li, Zhao Song
Abstract: Deep neural networks (DNNs) have demonstrated dominating performance in many fields, e.g., computer vision, natural language progressing, and robotics. Since AlexNet, the neural networks used in practice are going wider and deeper. On the theoretical side, a long line of works have been focusing on why we can train neural networks when there is only one hidden layer. The theory of multi-layer neur... ▽ More
Submitted 9 November, 2018; **originally announced** November 2018.

Step 3-2: Scrap the title, author, abstract, subject, and meta information by visiting the urls collected in Step 3-1 (You will earn extra credits if you collect the data from different parts of HTML source rather than from the same parts as in the provided script)

Step 4: Export the collected data with the file name "Deep learning arXiv papers.csv"