Example 1: Price forecast and virtual bidding of electricity contracts in the wholesale market

Great volatility exists in the wholesale electricity prices. With increasing penetration of renewable generation, such volatility could be magnified given the intermittent nature of the renewable energy sources. This project aims to develop efficient learning schemes that predict the real time electricity markets based on day-ahead information. Data can be collected from http://oasis.caiso.com/mrioasis/logon.do. Features such as historical day-ahead prices and load forecasts from PJM interconnection will be used for the prediction. The prediction would then be used for developing strategies of portfolio management in virtual bidding (see e.g. https://en.wikipedia.org/wiki/Virtual_bidding).

Example 2: Extraction road networks from satellite imagery

Recent flooding in Bangladesh and Hurricane Maria in Puerto Rico observe the need for accurate mapping during the disaster response. However, most existing map features such as roads, building footprints, and points of interest are primarily created through manual techniques. This project aims to develop advanced computer vision algorithms to process remote sensing data from https://spacenetchallenge.github.io/Competitions/competitionsSummary.html and automatically extract features to map the road networks from satellite imagery.

Example 3: Automatic music soundtrack generation

In order to augment the creative process, many tools now exist for helping artists write music. However, computer music generators typically operate at the note level and are able to capture long-term dependencies. In addition, they lack the expressive richness and nuance of performed music. This project aims to develop a novel deep learning model trained on the music dataset from http://www.piano-midi.de/ to compose music that reflect the nuance features such as pitch interval. timbre, tempo, rhythm.

Example 4: Fake Yelp review detection

Online reviews have become an important factor when people make purchase and business decisions. The increasing popularity of online reviews also stimulates the business of fake review writing, which refers to paid human writers producing deceptive reviews to influence readers' opinions. This project aims to building a classifier that takes the review text and the basic information of its reviewer as input (https://www.yelp.com/dataset/challenge) and outputs whether the review is reliable.

Notes:

- **Ideas**: Two of the main machine learning conferences are ICML and NIPS. You can find papers from recent ICML conferences online: http://icml.cc All NIPS papers are online, at http://books.nips.cc/. You can browse some of the recent machine learning papers to get inspired.
- **Datasets:** Kaggle, UCI Machine Learning Repository, Amazon AWS Open Data, Google Dataset Search
- **Example projects**: Stanford machine learning course projects http://cs229.stanford.edu/proj2017/