Hume-Nash Machines
Context-Aware Models of Learning and Recognition

Rocco Tripodi, Marcello Pelillo
ECLT – European Centre for Living Technology - Ca' Foscari University of Venice

Introduction

With the spectacular growth and the ubiquity of network data, classical (feature-based) approaches to machine learning and pattern recognition are no longer viable.

More sophisticated “context-aware” approaches, which exploit local environmental information, are needed in order for next-generation machines to cope with the increasing complexity of real-world applications.

In particular, this is of utmost importance in real-world scenarios involving several agents interacting in a complex environment using multiple cues and modalities.

Methodology

We propose Hume-Nash machines, a context-aware classification model based on:

- The use of similarity principles which go back to the work of British philosopher David Hume (“assign similar labels to similar objects”)
- Game-theoretic analysis introduced by Nobel laureate John Nash (“no incentive to unilaterally deviate from equilibrium”)

The intuition is to view learning problems as non-cooperative games, whereby the competition between the hypotheses of class membership is driven by contextual and similarity information encoded as payoff functions [1].

According to this perspective, the focus will shift from optimality of functions to equilibrium of (non-cooperative) games.

We modeled:
- The words to be disambiguated as the players of the games
- The senses of the words as the strategies that the players can choose
- The interactions among the players as a word similarity graph
- The payoff function as a sense similarity function

We used the replicator dynamics equation to compute the Nash equilibrium of the games.

In this way, it is possible to maintain the textual coherence associating each word to the most appropriate sense according to the senses that other words in the text are choosing.

The system was validated against unsupervised, semi-supervised and supervised algorithms.

We used four datasets for the WSD task and on each of them we obtained higher or comparable results compared to state-of-the-art systems. For the Entity Linking task we validated our system on two datasets and also on this task we obtained higher or comparable results.

Word Sense Disambiguation and Entity Linking

Word Sense Disambiguation is the task of identifying the intended meaning of a word based on the context in which it appears. It has been studied since the beginnings of Natural Language Processing and today it is still a central topic of this discipline. This is because it is important for many NLP tasks such as text understanding, text entailment, machine translation, opinion mining, sentiment analysis and information extraction.

All these applications can benefit from the disambiguation of ambiguous words, as a pre-processing step, in order to identify the words that are semantically relevant to the intended context.

In this way, it is possible to maintain the textual coherence associating each word to the most appropriate sense according to the senses that other words in the text are choosing.

The system was validated against unsupervised, semi-supervised and supervised algorithms.

We used four datasets for the WSD task and on each of them we obtained higher or comparable results compared to state-of-the-art systems. For the Entity Linking task we validated our system on two datasets and also on this task we obtained higher or comparable results.

Document Clustering Games as Dynamic Selection

Document clustering is a well-known task, that involves textual data. The objects to be clustered can have different characteristics, varying in length and content. Popular applications of document clustering aims at organizing news, novels, news documents and medical documents. It is a fundamental task in text mining, with different applications that span from document organization to language modeling. Clustering algorithms tailored for this task are based on generative models, graph models and matrix factorization techniques. A general problem, common to all these approaches, involves the temporal dimension. In fact, for these approaches it is difficult to deal with datasets that evolve over time and in real world applications documents are streamed continuously.

Furthermore, this problem can be more severe in case of huge datasets, because of scalability issues. With our approach we overcome these problems, in fact, we can classify new objects according to the information on previous clusterings. The problem of clustering new objects is defined as a game, in which we have labeled players (clusters) objects, which always play the strategy associated to the unmarked and unlabeled players which try to learn their strategy according to the strategy that their co-players are choosing.

The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects. The geometry of the data is modeled as a similarity graph, whose nodes are players (documents), and the games are played only on labeled objects.

The payoff of the games as opinion similarity is defined as:

\[ P = \sum_{i,j} Q_{ij} \delta_{ij} \]

where \( Q_{ij} \) is a similarity matrix of the objects. In this way, it is possible to maintain the influence that positive and negative players have on the opinions of the others, about a particular product, on the underlying social network.

References


Example Setup

Consider:

A set of objects \( R = \{r_1, ..., r_n\} \)

A set of labels \( L = \{l_1, ..., l_m\} \)

The goal is to label each object \( r_i \) with a label \( l_j \).

To this end, two sources of information are exploited:

- Local measurements which capture the salient features of each object viewed in isolation.
- Contextual and similarity information, expressed in terms of a payoff matrix of \( R = \{P_{ij}\} \).

Contact-Aware classification as a Non-Cooperative Game

We shall formulate the contact-aware classification problem as a non-cooperative game, where:

- Objects to be labeled \( r_i \)
- Class labels \( l_j \)
- Weighted labeling assignments \( A \)
- Contextual constraints \( C \)
- Payoff function \( R \)

The coefficient \( \alpha_{ij} \) measures the strength of compatibility between the two hypotheses: \( i \) is labeled \( k \), and \( j \) is labeled \( l \), or in other words the payoff that players get when player \( i \) plays strategy \( k \) and player \( j \) plays strategy \( l \).

According to this formulation, contact-aware classification = Nash equilibrium Payoff function:

\[ \min_{A, C} \sum_{r_i} \sum_{l_j} \alpha_{ij} \delta_{ij} \]

Where \( p \) is a probability distribution that denotes, for each player, the probability that it chooses a determined strategy.

System dynamics:

\[ \frac{dx}{dt} = \frac{1}{N} \sum_{j} \alpha_{ij} A_{ij} x_j \]

In this way better than average strategies grow at each iteration.

Contact Aware Nonnegative Matrix Factorization Clustering

We propose a method to refine the clustering results obtained with the nonnegative matrix factorization technique, imposing consistency constraints on the final labeling of the data.

The Dynamic of Opinion Diffusion in Social Networks

Social influence shapes every person's practices, judgments and beliefs is a truism to which anyone will readily assent. A child's master his native dialect down to the finest nuances; a member of a tribe accepts a common belief and tries to conform to it by imitating the other members. (Sorokin, R., 1990)...

The exponential growth in the popularity of the online social networks such as Facebook, Twitter has led to a lot of new research in understanding basic sociological phenomena such as opinion and consensus formation.

Psychologists discovered that people tend to adopt their opinions to their social environment trying to minimize the divergence among their opinions and the one of their friends. How the opinion on a particular topic can influence the opinions of other people in a social network?

Model:

- The players of the games as the users on the network.
- The strategies of the games as opinions
- The payoff of the games as opinion similarity
- The interactions among the players as a weighted similarity graph

This model will be tested on real world datasets to test its prediction power. The datasets will be constructed using the reviews posted by users on Amazon, Reddit and TripAdvisor. In this way it is possible to analyze the reviews and to measure the influence that positive and negative reviews have on the opinions of the others, about a particular product, on the underlying social network.