Project Proposal

Davide Berdin, Tobias Famulla
{davide.berdin.0110},{tobias.famulla.7003}@student.uu.se

Machine Learning
Department of Information Technology

April 13, 2015

1 Machine Learning for Date Matching

1.1 General Idea

Online Dating is omnipresent in our modern society. The amount of people looking for a potential partner or a fling is quite high and so is also the amount of suggested not fitting persons.

Tinder moved Online Dating into a new direction but still does not provide any automatic suggestions of whom could be a fitting match. To improve this solutions of matching automatically have already be written but different approaches have not been evaluated.

As our project in the subject "Machine Learning", we want to evaluate different approaches to find better matching possible partners using face recognition and analysis of mutual interests.

1.2 Machine Learning Parts

In our approach, we want to write a software, in which the user rates different person profiles including the picture and some properties like hobbies, interests, movies or books on a yes and no scale, whether the user would date the person or not. After a certain amount of cylces, the software starts to filter out persons, which seem unlikely to be a match, based on the face and properties. If the software works successfully, the "yes"-rate after a high amount of rounds will be significantly higher than the probability without the algorithm.

As the scientific part of the project, we want to evaluate different face recognition algorithms to find proper one for this application and also evaluate stategies to use machine learning algorithms to classify and optimize the users type for dating. Second we will evaluate algorithms to train the software on the properties.

We found already a software which uses Eigenfaces for the face recognition and attaches to Tinder. Our project should extend this approach and also test other algorithms like Fisherfaces, Local Binary Patterns Histogramms and explore differnt classifiers.

To ease the work, we will use different librarys, which are already published as open source software (see below)

1.3 Possible Evaluation

One problem of this project might be that we have a small data basis to evaluate if the algorithms are successful. One value to measure it, is to show the user profiles after a training phase, from whom the algorithm assumes the user will click on "yes" or "no" and evaluate if the algorithm is right significantly often. To check it for different types of people, we would try to encourage friends to use the software to summarize the data later.

1.4 Tools

We decided, that we would like to use the programming language Python, because a lot of libraries for machine learning are available and we could also used API to existing services like Tinder or Facebook to retrive real world profiles of people.

As general Machine Learning libraries, PyBrain and scikit-learn are available, which provide a variety of different Neuronal Network algorithms. For the Face Recognition part, we would look at OpenCV, facerec [1] and facereclib [2].

1.5 Prior Work

As mentioned above, a project called Tinderbox [3] is available to automatically match pictures of people and send them automatic matches. Additionally we found a variety of papers concerning face-recognition.

---

[1] https://github.com/bytefish/facerec
[2] https://github.com/idiap/facereclib
[3] https://github.com/crockpotveggies/tinderbox