

# Recruiting Analytics

The key task we are interested in is in matching people to jobs. The key entities here are (a) jobs, (b) job ads, (c) people, and (d) skills.

**Jobs** A job is indexed by its title, and has the following features: a list of skills, and amount of experience in these skills. Potentially, for instance as a pre-req., it could also have as feature other jobs, and amount of experience in those other jobs. The latter could also be generalized to a network/taxonomy over such jobs: this is available via governmental and other databases; e.g. Bureau of Labor Statistics.

**Job Ads** A job ad can also be thought of as a *variant* of a job as defined above. Thus, it is associated with a job title, and text comprising the job ad. The text consists of features associated with a job noted above: a list of skills, and amount of experience in these skills, other jobs and amount of experience in other jobs.

**People** A person has the following features: (a) a list of skills, a list of current/previous jobs, and duration/experience associated with previous jobs and skills; (b) education: school name, degree type; (c) other features?

There are also social networks connecting people, with weighted edges.

**Skills** These are features within the entities listed above, such as jobs, and people. There could also be unobserved/latent skills that we might be able to estimate given data.

**Our Approach** To start with, we can model this as a recommendation system problem, similar to the Netflix Challenge. So, we observe/create a partially observed matrix, with rows as people, and columns as jobs. The entries are *suitability* values that could be inferred, for instance from the experience of the person at that job. The task then, is to predict the suitability of people to jobs in entries which we have not observed.