ML for Predictive Resume Screening

Sept 17 2018

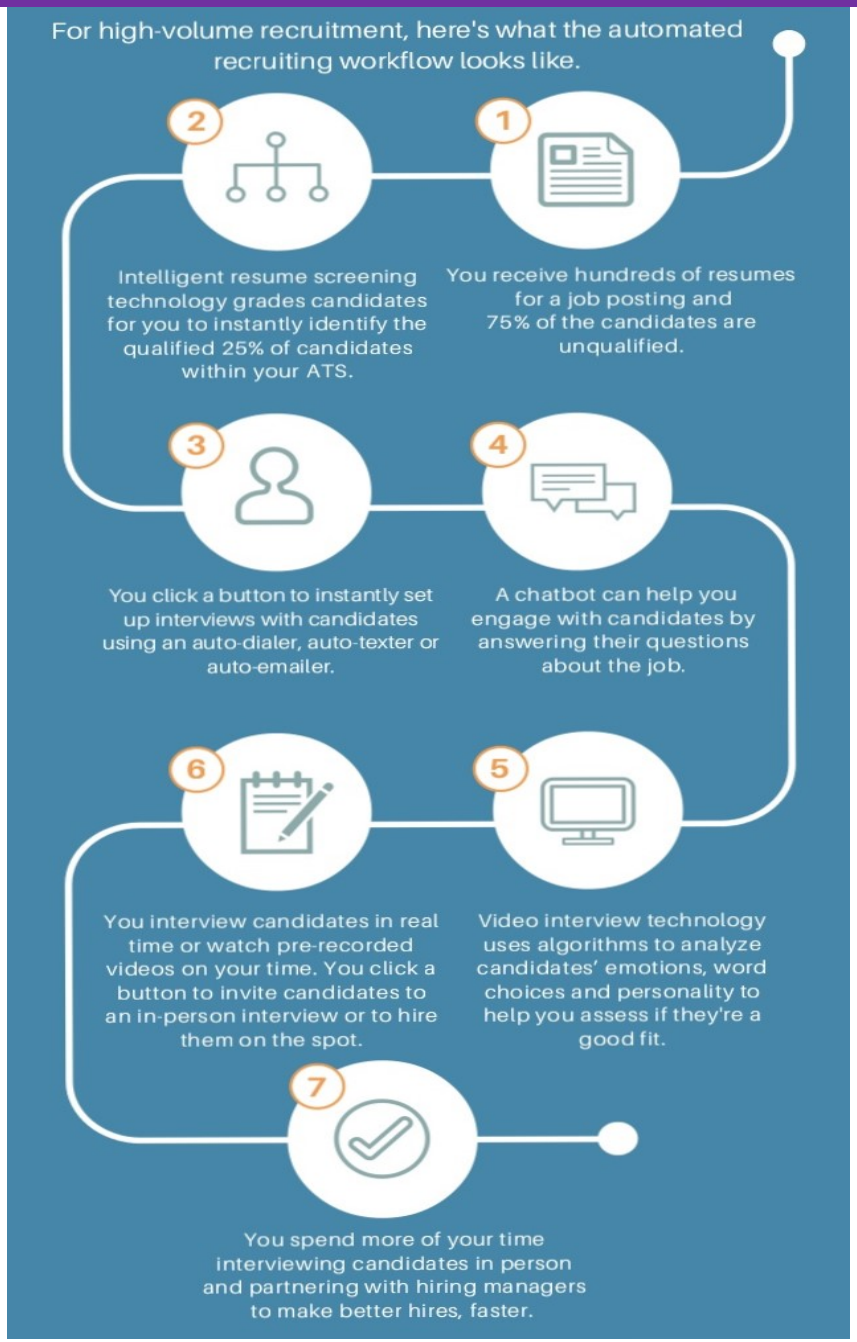
HIREGROUND HR Software & Solutions Overview

- For **each job has 67 applicants** in average, each job can have over 300 applicants easily
- **Each client has around 10,000 applicant** profiles
- Our system has more than **300,000 resumes for 4255 jobs**
- Average **2500 applicants per day**
- This number increasing daily ...
- We have enough data for AI purpose

Applicant Tracking System (ATS) Software

Automates Full Cycle Hiring

- Screening
- Manage Interviews
- Answering Questions
- On-boarding
- Archiving and Followup



(Source: ideal-recruit)

Resume Parsing

Richard Manns

Convert to text

Extract Sections (Work Experience, Education..)

NLP Unit, extract keywords, n-gram

Generate Feature Vector

NLP Extraction

Keywords

These are descriptive features about the observed and target text, including the length of the text, and indicator features such as whether or not the resume has a cover letter, or the count of experiences the applicant has.

N-Gram Word Intersection Based Features

N-grams for both observed and target text and calculate their intersection count and co-occurrence count as features.

N-Gram IR Scoring Features

We score each resume and job description using the usual TF, TFIDF and Okapi BM25 scores used for document ranking at the N-gram level and five statistical features are extracted

N-Gram LSI Based Features

Identify patterns in the relationships between terms and concepts contained in the text corpus.

Algorithms

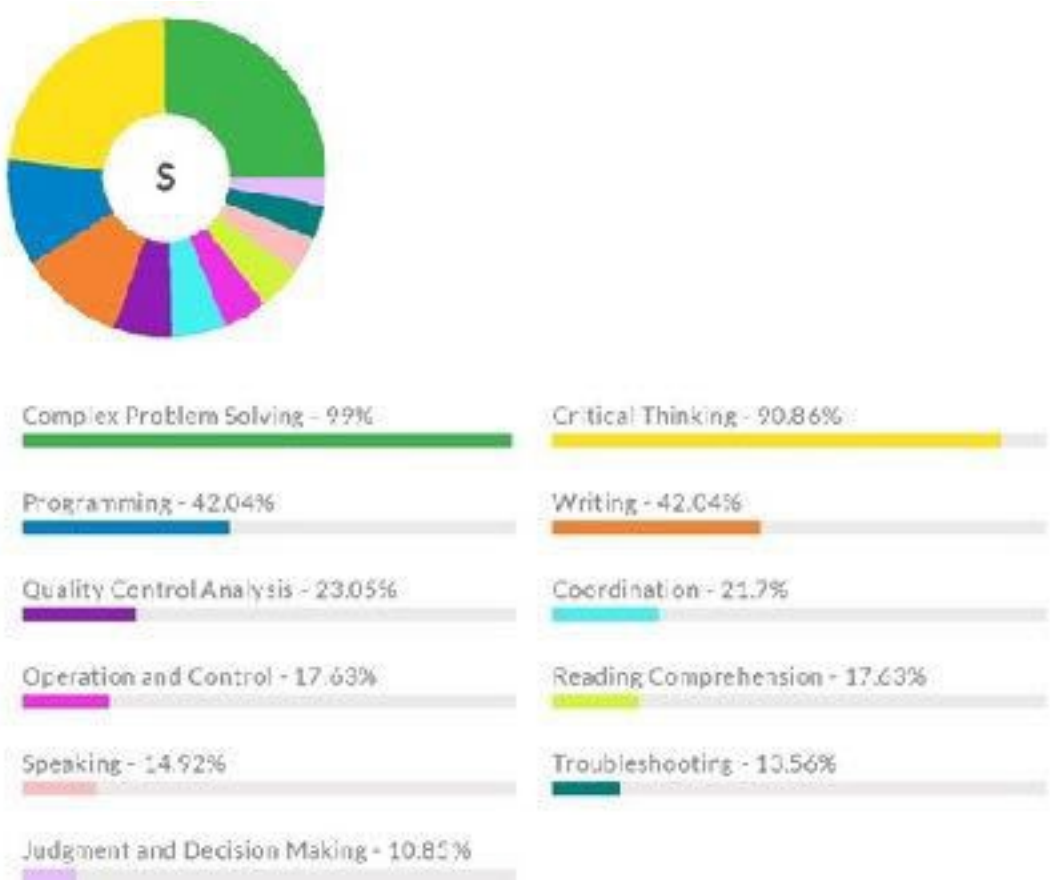
- SVM (Support Vector Machine)
- Stacked Neural Network

Screening


EXCELLENT
GOOD
MODERATE
AVERAGE
POOR
HireXtra AI Matching Algorithm Says Aravind Profile is Moderate


Required Skill Match
Preferred Skill Match
Position Match
N/A
Location Match
Qualification Match
Experience Match

Matching


S
Complex Problem Solving - 99%
Critical Thinking - 90.86%
Programming - 42.04%
Writing - 42.04%
Quality Control Analysis - 23.05%
Coordination - 21.7%
Operation and Control - 17.63%
Reading Comprehension - 17.63%
Speaking - 14.92%
Troubleshooting - 13.56%
Judgment and Decision Making - 10.85%

Inference


Thinking Creatively
99%
Interacting With Computers
75.57%
Analyzing Data or Information
50.66%
Documenting/Recording Information
34.91%
Inspecting Equipment, Structures, or Material
34.72%
Getting Information
32.54%
Judging the Qualities of Things...
32.24%
Drafting, Laying Out, and Specifying...
29.68%
Provide Consultation and Advice to...
29.6%
Monitor Processes, Materials, or Surroundings
29.12%
Organizing, Planning, and Prioritizing Work
28.49%
Guiding, Directing, and Motivating Subordinates
27.54%

TO-DO LIST FOR CURRENT ALGORITHM
Project Scope

1. All of our historical data will be trained
   - This needs to use of our current Postgres DB

2. Each section of Resumes/ each category is needed to built as separate models
   - For example: education, work experience, contact..

3. Visualization
   - Needs knowledge of BackEnd/Frontend coding
   - Ruby/Angular

4. Apply the algorithm based on deep learning collaborated with Uof A

Tools

- Postgres (PSQL)
- Python
- Machine Learning Basics
- Ruby (optional)
- Angular (especially functions related to visualization, such as chart.js)
- Data analytical tools. (e.g: google analytics, google charts, R)
- Regular Expressions
- Perl (optional)
- MongoDB (NoSQL) (optional)