

Informe

Ricardo Arango Giraldo

14 de mayo de 2018

La base de datos *producto-prueba* consta de las siguientes tablas:

```
## Tables_in_producto-prueba
## 1      ar_internal_metadata
## 2      candidate_applications
## 3      candidate_statistics
## 4      candidates
## 5      employer_statistics
## 6      employers
## 7      job_applications
## 8      jobs
## 9      schema_migrations
## 10     work_experiences
```

en las cuales se almacena información de **empleadore, trabajos y candidatos**, así como de **las condiciones laborales** de las mismas.

1. ¿Cómo definir la calidad del candidato al postularse a un trabajo?

Hay que ver si las características del candidato se ajustan a los requisitos que tiene el trabajo. El principal requisito presente en los datos es la experiencia.

```
##
##      1-2 Years      2-3 Years      +3 Years Less than 1 Year
##      1421          605          1335          1540
```

Por ende, los candidatos mejor calificados son los que tienen más experiencia (+3 años). Ya otra característica que se puede tener en cuenta es la del cargo que actualmente ocupa, o el último que ocupó.

2. ¿Qué porcentaje de aplicaciones semanales son de candidatos bien calificados?

Primero que todo tengamos en cuenta que una misma persona puede ser un candidato a diferentes trabajos. Siendo así, los idUser no necesariamente deben ser únicos para este caso particular.

Tenemos un total de 14929 candidatos, de los cuales 1326 son candidatos con la mejor calificación. Lo que quiere decir que un 8.8820417% son candidatos muy bien calificados.

Ahora bien, lo anterior es en general así que analicemos el porcentaje promedio por semana: con un total de 29 semanas, tenemos un promedio semanal de 0.3062773% candidatos excelentemente calificados que aplicaron a algún trabajo en el lapso de esas 29 semanas.

3. ¿Cómo abordarías este problema? Cualquier idea sobre qué mecanismos específicos podría implementarse para hacer que nuestra comparación sea lo más eficiente posible?

Una regresión logística donde la variable respuesta sería el evento contratado. Para esto es necesario más datos para entrenar el modelo, así como más variables para enriquecer el análisis.

Completar el perfil de los candidatos con su trayectoria académica, así como indicar qué tipo de conocimientos son necesarios para ejercer el cargo.