

Job Recommendation System

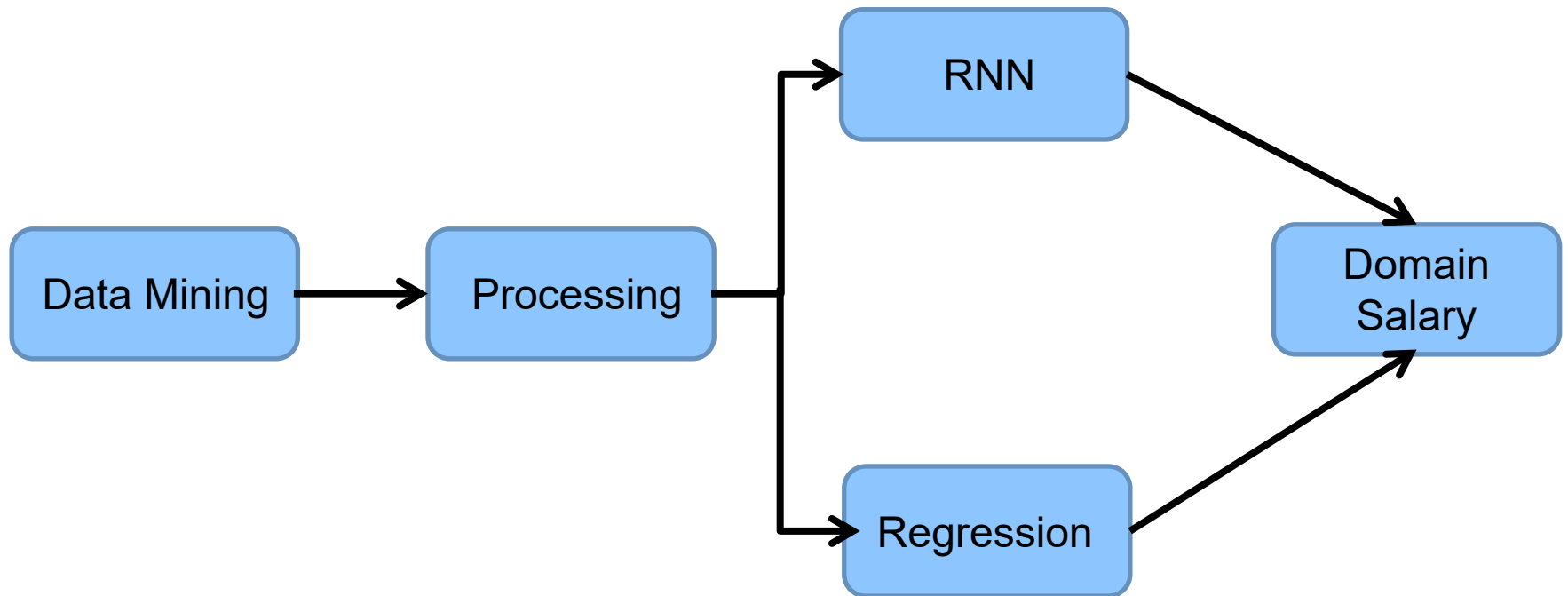
yc3993 Yutong Chen

yz4131 Yuqin Zhao

sr3849 Shuai Ren



Workflow



1.Data Mining

TRY:

LinkedIn API



Craw from LinkedIn Comment

| | Name | E-mails | Headline |
|---|---------------------|---------------------------------|---|
| 0 | Vaishvik Kumar | kvaishvik24@gmail.com | Software Engineer @ Samsung R&D Institute,Noida |
| 1 | Rutuja L. | ru2jalimaye11@gmail.com | Associate BI |
| 2 | Vedant Sharma | vedant.12sharma@gmail.com | Student at PES University |
| 3 | Pavan Kumar pothuru | pavankumar2302@gmail.com | Machine learning Enthusiast, Graduated from In... |
| 4 | Shikhar Tiwari | shikhartiware.bme18@itbhu.ac.in | Analyst Intern @RoaDo Ex- Intern @IIM Ahmeda... |
| 5 | Mohithsingh M | mohithsingh7@gmail.com | Test Lead at HCL Technologies Ltd |
| 6 | Ashwini Deshmukh | aashu.deshmukh06@gmail.com | Senior Analyst at Tiger Analytics Ex TCSer |
| 7 | Kirti Gupta | kirtigupta10007@gmail.com | Junior Data Analyst at Indian Navy |
| 8 | Kireina Christensen | kireinachristensen@gmail.com | IT Support Technician at Clarity Technology Gr... |
| 9 | Raghav Khandelwal | raghavkhandelwal2710@gmail.com | Upcoming Summer Intern '22 @ Microsoft Micro... |

1.Data Mining

Data World: <https://data.world/datasets/jobs>

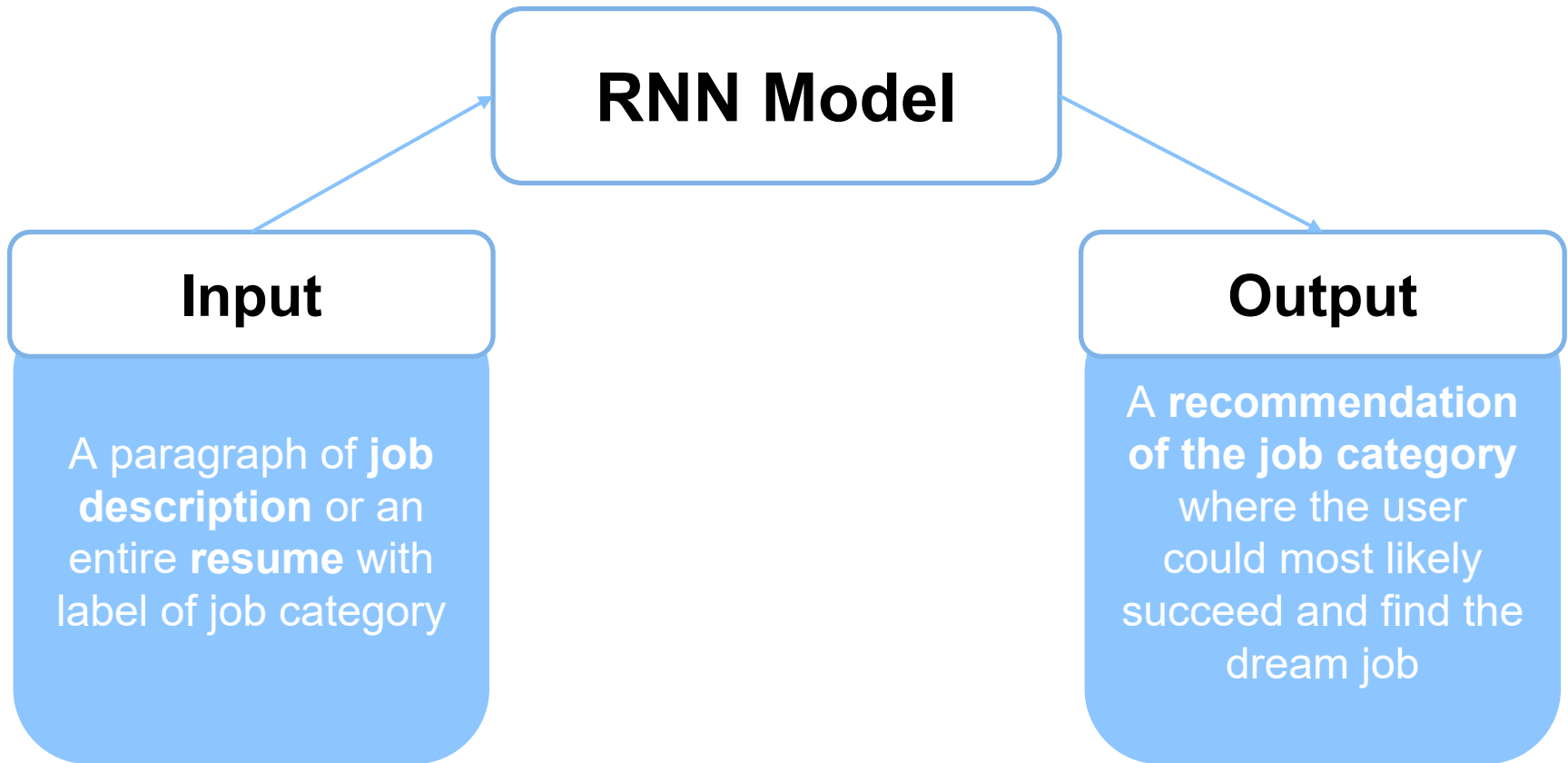
NYC Open Data: <https://data.cityofnewyork.us>

human resources leader 18 years experience indian global markets across analytics banking andtechno hr
proactive resultoriented hr professional experienced talent acquisition hrbpsales demonstrated historychr
i undertaken global business projects spanning across 3 continents set regional offices across india hire hr
an hr professional 4 years experience keen interest designing processes analytics i bring onboard expe hr
experienced human resources demonstrated history working human resources industry skilled inoperathr
12 years experience ir hr skills ability institutionalising industrial relations organisation culture asper busihr
working hr demonstrated history working education management industry skilled humanresources ma hr
experienced senior process executive demonstrated history working banking services industrystrong o hr
human resource professional experience across strategy business partnering large scale mnacs known fchr
experienced talent development and human resources executive demonstrated history working varied hr
reporting directly chief corporate development officer i responsible development andimplementation chr
hr professional 7 years experience multiple hr functions talent acquisition hr analyticsand hr business pchr
digital online marketing grow promote develop business wheezine best digital marketing companyoffe designing
experienced fashion designer demonstrated history working apparel fashion industry skilled intrend an designing
we strong team experienced dedicated web designersweb developers animation expertsnetworkingen designing
rishikesh cofounder principal architect multidisciplinary design practice design nonstop nashikbelieving designing
as interior designer reate aesthetically pleasing environments represent clients lifestyle taste lan desidesigning
experienced junior architect demonstrated history working architecture planning industry skilled inauto designing
jrarchitect skyline associates see designing

| jobType | degree | major | industry | yearsExperience | milesFromMetropolis | salary |
|----------------|-------------|------------|-----------|-----------------|---------------------|--------|
| CFO | MASTERS | MATH | HEALTH | 10 | 83 | 130 |
| CEO | HIGH_SCHOOL | NONE | WEB | 3 | 73 | 101 |
| VICE_PRESIDENT | DOCTORAL | PHYSICS | HEALTH | 10 | 38 | 137 |
| MANAGER | DOCTORAL | CHEMISTRY | AUTO | 8 | 17 | 142 |
| VICE_PRESIDENT | BACHELORS | PHYSICS | FINANCE | 8 | 16 | 163 |
| MANAGER | DOCTORAL | COMPSCI | FINANCE | 2 | 31 | 113 |
| CFO | NONE | NONE | HEALTH | 23 | 24 | 178 |
| JUNIOR | BACHELORS | CHEMISTRY | EDUCATION | 9 | 70 | 73 |
| JANITOR | HIGH_SCHOOL | NONE | EDUCATION | 1 | 54 | 31 |
| VICE_PRESIDENT | BACHELORS | CHEMISTRY | AUTO | 17 | 68 | 104 |
| JANITOR | HIGH_SCHOOL | NONE | HEALTH | 24 | 30 | 102 |
| CEO | MASTERS | PHYSICS | EDUCATION | 7 | 79 | 144 |
| JUNIOR | NONE | NONE | OIL | 8 | 29 | 79 |
| JUNIOR | MASTERS | MATH | FINANCE | 21 | 26 | 193 |
| JANITOR | NONE | NONE | AUTO | 21 | 81 | 47 |
| CTO | MASTERS | BIOLOGY | SERVICE | 13 | 8 | 172 |
| JUNIOR | MASTERS | PHYSICS | SERVICE | 1 | 91 | 47 |
| VICE_PRESIDENT | MASTERS | LITERATURE | SERVICE | 23 | 43 | 126 |
| CEO | MASTERS | LITERATURE | SERVICE | 23 | 66 | 122 |
| CEO | MASTERS | PHYSICS | EDUCATION | 9 | 99 | 95 |

≈ 100,000 Data

2. RNN Model



2. RNN Model

Data preprocessing

- Changed label to number
- Deleted meaningless words
- Word embedding

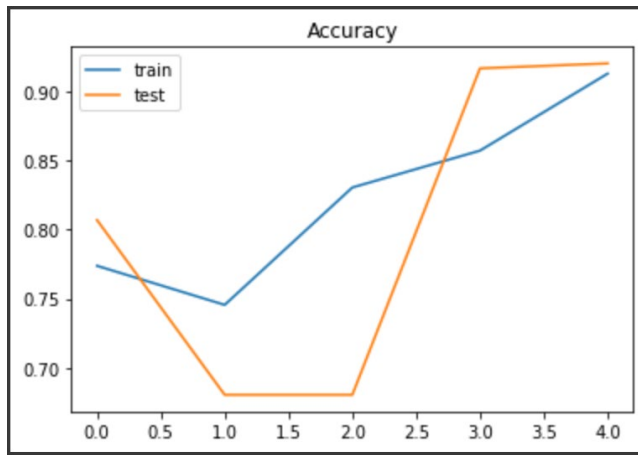
Building model

Due to the limitation by the GPU and time, we chose to build shallow RNN model

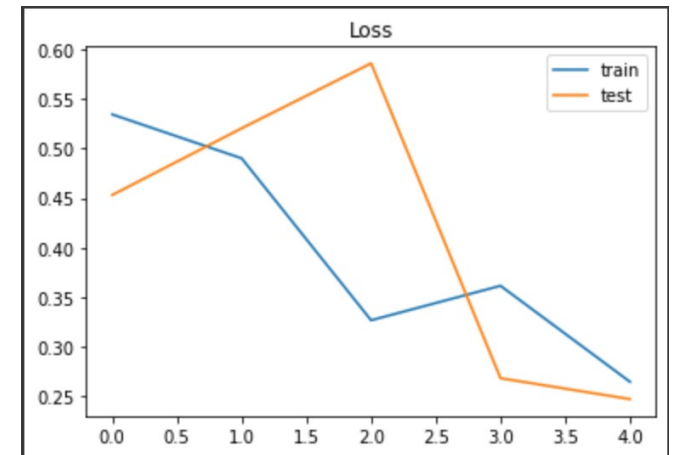
| Layer (type) | Output Shape | Param # |
|---------------------------|------------------|---------|
| embedding_4 (Embedding) | (None, 250, 100) | 50000 |
| lstm_8 (LSTM) | (None, 250, 50) | 30200 |
| lstm_9 (LSTM) | (None, 50) | 20200 |
| dense_4 (Dense) | (None, 27) | 1377 |
| Total params: 101,777 | | |
| Trainable params: 101,777 | | |
| Non-trainable params: 0 | | |

2. RNN Model

Training and Result



The validation accuracy reached more than 90%



Testing lost drop to 10%

```
predict('Making the right decision is a difficult but interesting thing, you must calculate the outcome of each choice and calculate the impact of each outcome. After all, I am a  
'data scientist')
```

```
predict('Currently I am mainly working on back-end, such as micro-service, automated workflow and event processing pipeline, but I also have experience in deploying  
'software design')
```

3. Salary Prediction

F
E
A
T
U
R
E
S

Job Type

Degree

Major

Industry

Working Experience

Location(Miles from Big Cities)

R
E
G
R
E
S
S
I
O
N

3. Salary Prediction

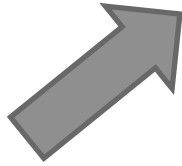
| label | features | prediction |
|-------|----------------------------|--------------------|
| 19 | (27, [4, 7, 11, 22, 25...] | 21.053326297405064 |
| 20 | (27, [4, 7, 11, 22, 25...] | 16.74219323944338 |
| 20 | (27, [4, 7, 11, 22, 25...] | 20.272407349510146 |
| 20 | (27, [4, 7, 11, 22, 26...] | 19.849790186483375 |
| 20 | (27, [4, 7, 11, 22, 26...] | 19.459330712535916 |
| 20 | (27, [4, 7, 11, 22, 26...] | 15.55473597306134 |
| 20 | (27, [4, 7, 11, 22, 26...] | 15.164276499113896 |
| 20 | (27, [4, 8, 11, 22, 25...] | 13.491686896564715 |
| 20 | (27, [4, 8, 11, 22, 26...] | 18.55158121334199 |
| 20 | (27, [4, 8, 11, 22, 26...] | 14.256526999919956 |
| 21 | (27, [4, 7, 11, 22, 25...] | 25.33230166628742 |
| 21 | (27, [4, 7, 11, 22, 25...] | 21.42770692681286 |
| 21 | (27, [4, 7, 11, 22, 25...] | 20.646787978917942 |
| 21 | (27, [4, 7, 11, 22, 25...] | 20.256328504970483 |
| 21 | (27, [4, 7, 11, 22, 25...] | 18.304031135233217 |
| 21 | (27, [4, 7, 11, 22, 26...] | 29.611277035169778 |
| 21 | (27, [4, 7, 11, 22, 26...] | 14.773817025166437 |
| 21 | (27, [4, 8, 11, 22, 26...] | 21.28479753097419 |
| 21 | (27, [4, 8, 11, 22, 26...] | 18.16112173939453 |
| 22 | (27, [4, 7, 11, 22, 25...] | 26.503680088129798 |

LR Test set RMSE = 19.7213080340121

3. Salary Prediction

HTML User Interface

`pickle` — Python object serialization



REGRESSION

MODEL



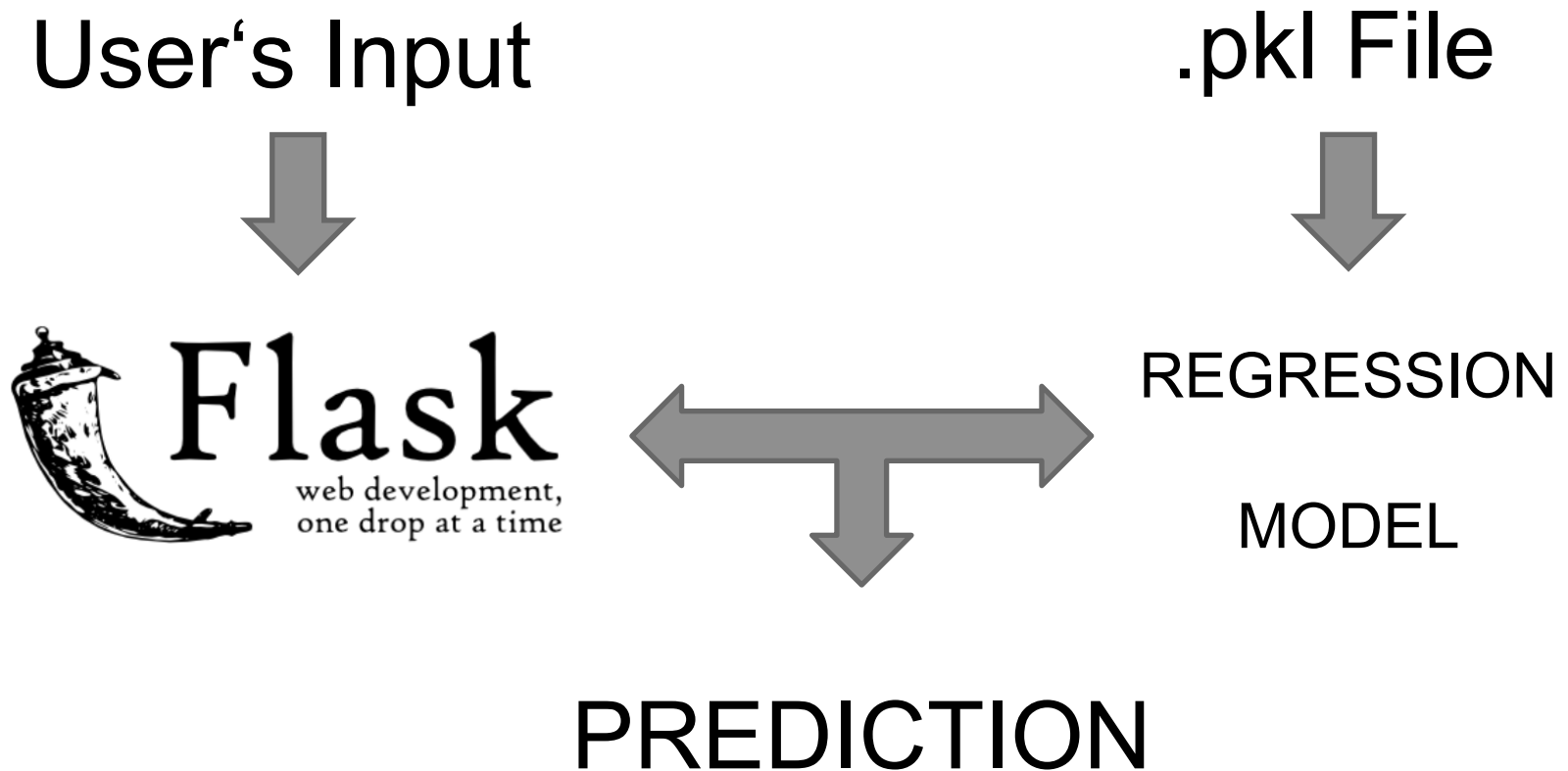
.pkl File



HTML

3. Salary Prediction

HTML User Interface



3. Salary Prediction

HTML User Interface

<https://salarypredic-g7.herokuapp.com/>

Input values:

jobcat: JUNIOR

degree: MASTERS

major: COMPSCI

industry: WEB

workexperience: 0

dis: 0

Submit

jobcat : JUNIOR degree : MASTERS major : COMPSCI industry : WEB workexperience : 0 dis : 0

Salary Predicted:

106k

#Data from PayScale shows that computer science professionals with a master's degree, specifically, see an average salary of \$101,000

4.Drawbacks

1. UI for Deep Learning Model
2. Better way for Data Mining

Thank You For Listening !