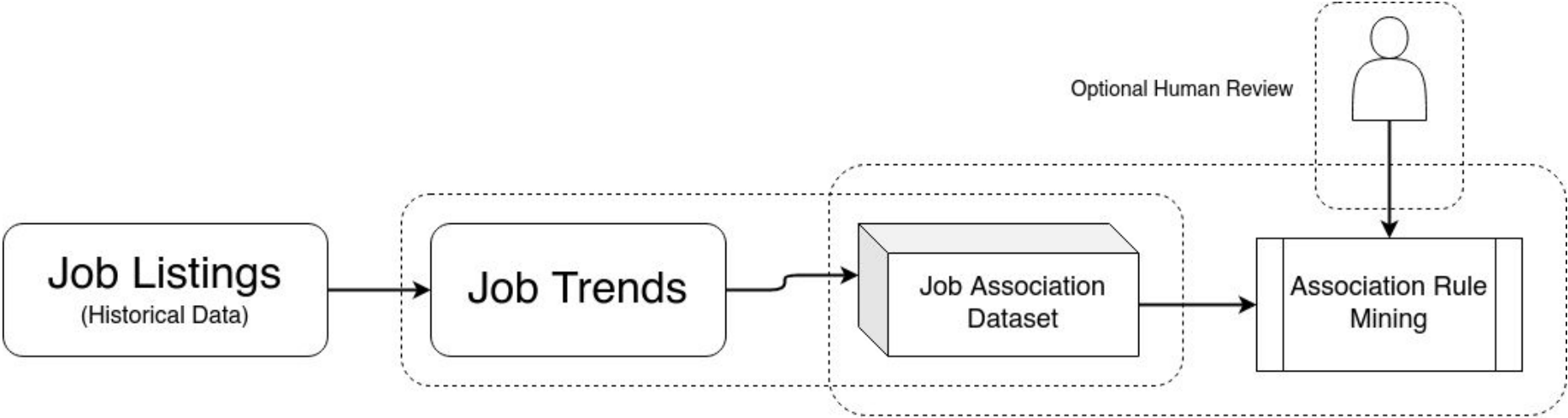
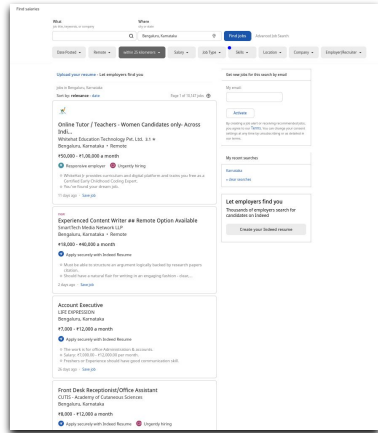


# ML Pipeline



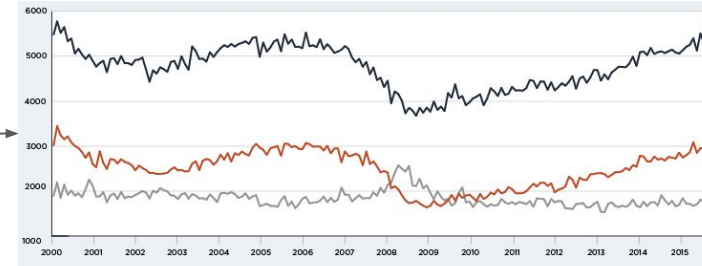
# Data collection & Dataset



Job site listings



Parsed



Job Trends

# Bayesian Association Rule Mining

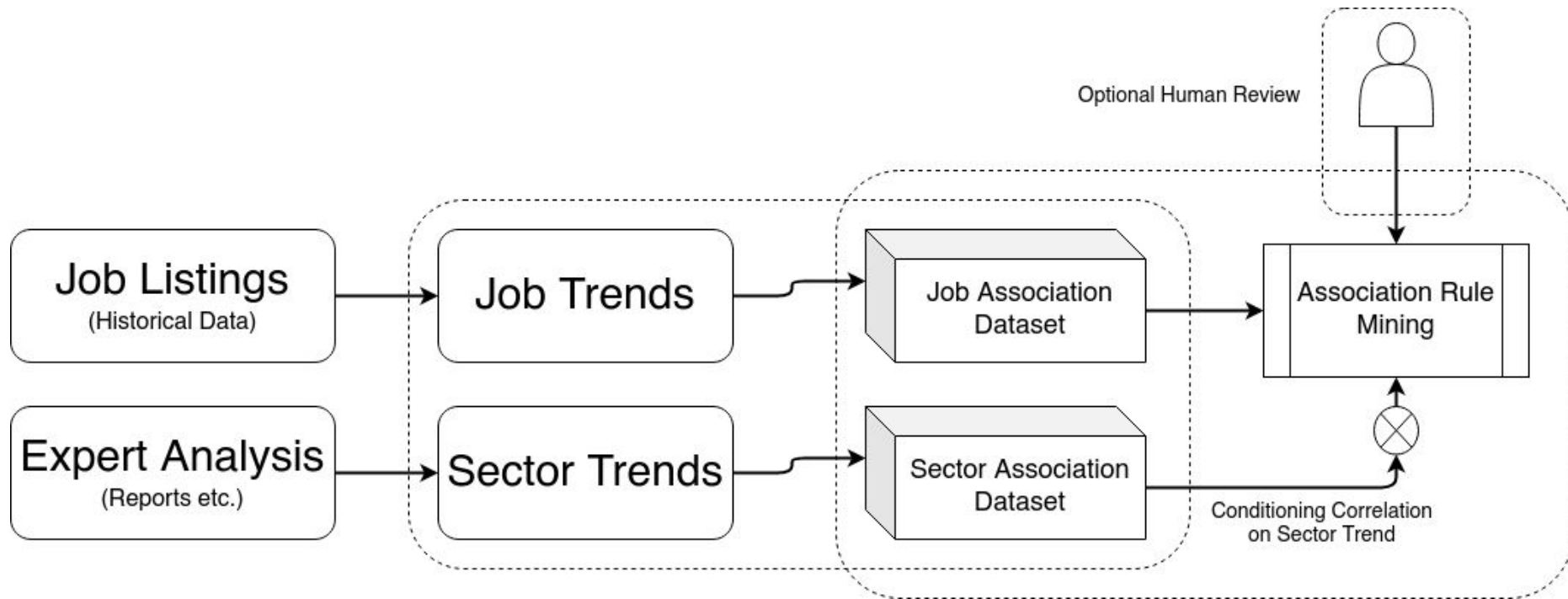
Naive Bayes: 
$$p(C_k | \mathbf{x}) = \frac{p(C_k) p(\mathbf{x} | C_k)}{p(\mathbf{x})}$$

Conditioning job positions:  $P(\text{job}_i | \text{job}_j) = \frac{P(\text{job}_i, \text{job}_j)}{P(\text{job}_j)}$

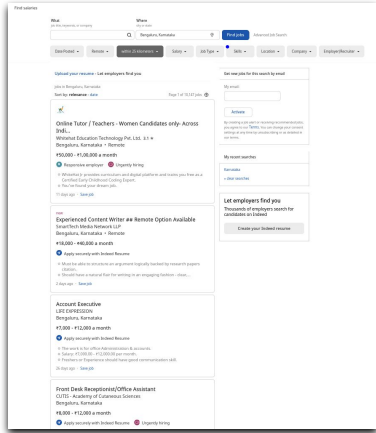
## Result & Analysis



## ML Pipeline - Modified



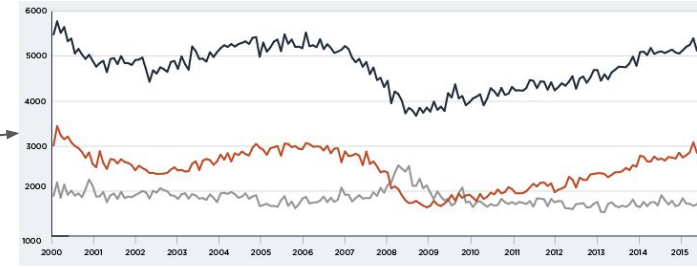
# Data collection & Dataset



Job site listings

```
{
  "job-title": "ACCOUNTING AND REPORTING R2R IFRS General Accounting US GAAP",
  "company": "Taggs",
  "date-posted": "11/11/2015",
  "location": "Bengaluru / Bangalore",
  "duration": "1-7 years",
  "salary": "350000 - 550000 INR",
  "description": "It's a Very Good Opportunity for the CA Inter Qualified, who all are look",
  "skills": "Skills : IFRS"
},
{
  "job-title": "Phone Banking Officer",
  "company": "Coppergate Consultants",
  "date-posted": "11/11/2015",
  "location": "Delhi",
  "duration": "Not Specified",
  "salary": "null",
  "description": "To resolve customer's queries on phone <br>",
  "skills": "Skills : Banking"
},
{
  "job-title": "US Accounting",
  "company": "Schema Technologies",
  "date-posted": "11/11/2015",
  "location": "Hyderabad / Secunderabad",
  "duration": "10-2 years",
  "salary": "Not Specified",
  "description": "US Accounting",
  "skills": "Skills : Accounting"
},
{
  "job-title": "Financial Auditors and Accountants - Canada",
  "company": "A V Immigration And Careers Consultancy Private Limited",
  "date-posted": "11/11/2015",
  "location": "Australia",
  "duration": "13-8 years",
  "salary": "1800000 - 3600000 INR",
  "description": "Financial auditors examine and analyze the accounting and financial recor",
  "skills": "Skills : Financial Audits"
}
```

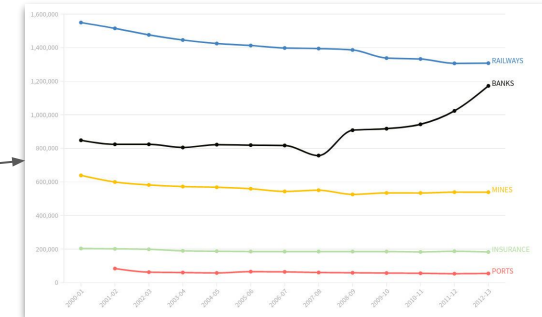
Parsed



Job Trends

YEAR	RAILWAYS	MINES	PORTS	INSURANCE	BANKS
2000-01	1549385	638741	203200.0	848000	824000
2001-02	1514816	599301	83771	201100.0	824000
2002-03	1475884	581880	62244	198400.0	824000
2003-04	1445428	572327	59941	189100.0	805000
2004-05	1424389	567911	57492	186800.0	822000
2005-06	1412434	559095	65347	185000.0	819000
2006-07	1397610	543009	63903	184700.0	817000
2007-08	1394418	550390	60211	185000.0	757000
2008-09	1386123	525501	58556	184900.0	908000
2009-10	1337533	534077	56878	185100.0	917000
2010-11	1332204	533834	55472	182700.0	943000
2011-12	1306210	538812	52934	186567.0	1023000
2012-13	1307253	538625	54390	182411.0	1172000
2013-14	1310818	543274	51409		1125450
2014-15	1326370	538945	44843		1154490
2015-16					

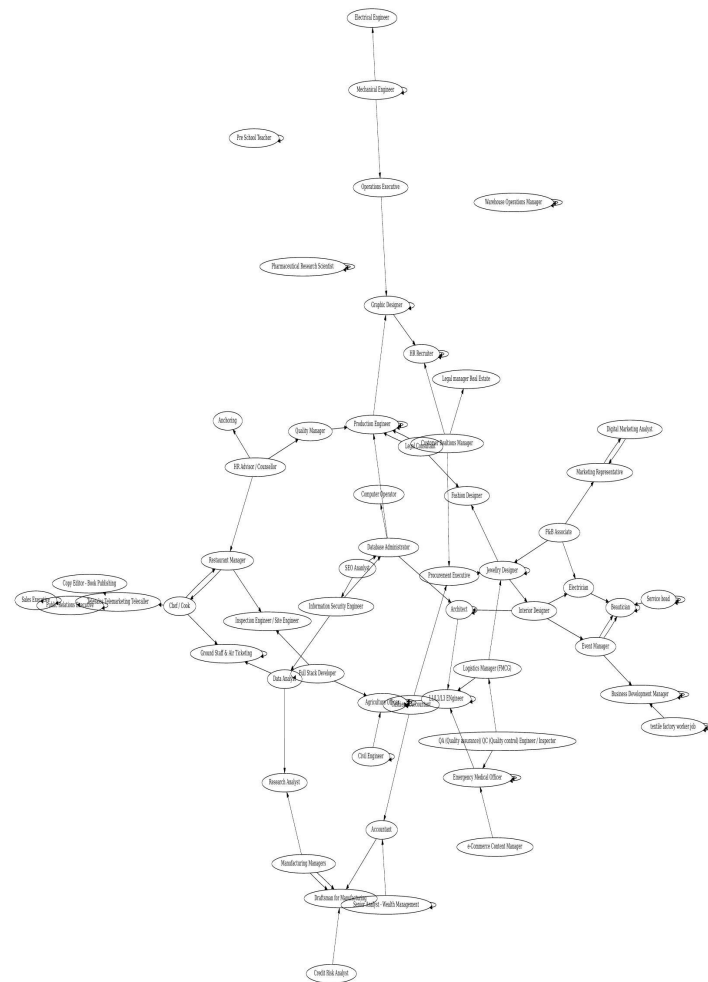
40716



Sector Trends

Reports etc.

## Results & Analysis



Using both sector and job position data:  $P(\text{job}_i | \text{job}_j) \cdot P(\text{sector}_i | \text{sector}_j)$