Big Data for Labour Market Intelligence

Capacity development programme 2022

Module 1: Technical training

Session 6
Data classification, data retrieval, practical exercise and coding hands-on

Speaker: Simone Perego

09/11/2022
Data Classification

• **Goal:**
  • Extract and structure information from data, to be provided to the presentation layer

• **Challenges:**
  • Handle massive amount of heterogeneous data written in different languages

• **Approach:**
  • Develop an adaptable framework, language dependent, tailored on different information features. Some relevant challenges:
    • Occupation feature classification: combined methods such as Machine Learning, Topic Modeling and Unsupervised Learning
    • Skill feature classification: another different combined methods, such as Text Analysis with corpus based or Knowledge based similarity

• **Features:**
  • Guarantee Explainable information extraction, logging classification methods and relevant features.
Junior Software Developer

As Junior Software Developer, you will develop excellent software for use in field mapping, data collection, sensor networks, street navigation, and more. You will collaborate with other programmers and developers to autonomously design and implement high-quality web-based applications, restful API's, and third party integration.

We’re looking for a passionate, committed developer that is able to solve and articulate complex problems with application design, development and user experiences. The position is based in our offices in Harwell, United Kingdom.

2512 – Software Developer

Skills: develop software, implement web based applications, problem solving, develop user experiences

Harwell, UK
Occupation pipeline

1. Language Detector
2. Pre Processing
3. Ontology based models
4. Ontologies
5. Machine learning model
6. Machine learning classifier
7. Classified items
Occupation Output

- Professional services managers
  - Manufacturing mining, construction, and distribution managers
  - Production and specialised services managers
    - Production managers in agriculture, forestry and fisheries
    - Information and communications technology service managers
  - Business services and administration managers
  - Administrative and commercial managers
    - Sales, marketing and development managers
    - Chief executives, senior officials and legislators
    - Managing directors and chief executives
    - Hotel and restaurant managers
  - Hospitality, retail and other services managers
    - Other services managers
    - Retail and wholesale trade managers
  - Managers
    - Supply, distribution and related managers
    - Agricultural and forestry production managers
    - Aquaculture and fisheries production managers
    - Information and communications technology service managers
    - Finance managers
    - Human resource managers
    - Policy and planning managers
    - Business services and administration managers not elsewhere classified
    - Sales and marketing managers
    - Advertising and public relations managers
    - Research and development managers
    - Legislators
    - Senior government officials
    - Traditional chiefs and heads of village
    - Senior officials of special-interest organizations
    - Managing directors and chief executives
    - Hotel managers
    - Restaurant managers
    - Sports, recreation and cultural centre managers
    - Services managers not elsewhere classified
    - Retail and wholesale trade managers
    - Child care services managers
    - Health services managers
    - Aged care services managers
    - Social welfare managers
    - Education managers
    - Financial and insurance services branch managers
    - Professional services managers not elsewhere classified
    - Manufacturing managers
    - Mining managers
    - Construction managers

Job Location Recognition

C#, XML and Data Software Developer
Scottish Government 4.5 stars 40 reviews - Edinburgh
£28,891 - £33,094 a year - Full-time, Part-time, Permanent

For the full job description and to apply, please visit the Work for Scotland website
Benefits
Pensions
Civil Service pension schemes may be available for successful candidates.

Apply On Company Site

Place Match

Edinburgh

NUTS 2021 + other sources

<table>
<thead>
<tr>
<th>Key</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>{ id : UKM25 }</code></td>
<td>{ 17 fields }</td>
</tr>
<tr>
<td><code>lau_1</code></td>
<td>UKM25</td>
</tr>
<tr>
<td><code>lau_2</code></td>
<td>Edinburgh, City of</td>
</tr>
<tr>
<td><code>idLau_1</code></td>
<td>Eastern Scotland</td>
</tr>
<tr>
<td><code>idLau_2</code></td>
<td>SCOTLAND</td>
</tr>
<tr>
<td><code>nutLiv_3</code></td>
<td>UNITED KINGDOM</td>
</tr>
<tr>
<td><code>nutLiv_0</code></td>
<td>UKM25</td>
</tr>
<tr>
<td><code>idNutLiv_2</code></td>
<td>UKM2</td>
</tr>
<tr>
<td><code>idNutLiv_1</code></td>
<td>UKM</td>
</tr>
<tr>
<td><code>idNutLiv_0</code></td>
<td>UK</td>
</tr>
<tr>
<td><code>name</code></td>
<td>Edinburgh, City of</td>
</tr>
<tr>
<td><code>country</code></td>
<td>UK</td>
</tr>
<tr>
<td><code>nutsLevel</code></td>
<td>3.0</td>
</tr>
<tr>
<td><code>idIso31161</code></td>
<td>UG</td>
</tr>
</tbody>
</table>
Job Location pipeline

- Language Detector
- Pre Processing
- Ontology based models
- Classified items
- Ontologies
Skill pipeline

Ontologies

Language Detector → Pre Processing → Ontology based models → Classified items
Fabasoft International Services GmbH

Application Developer (w / m)

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries
- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- Interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0
We offer you
<working>
<highly motivated team>
<professional>
<Certified Project Manager>
<Scrum>
<performance-based>

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries
- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- Interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0
You are the right one if:

- you are interested in designing, implementing and integrating specialized applications and you like to work with the development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries
- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- Interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0
You are the right one if

- you are interested in designing, implementing and integrating specialized applications and you like to work with cutting-edge development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products /SDKs as well as libraries
- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- Interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fasboosoft.com
Tel: +43 732 606162-0
You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertise use standard products / SDKs as well as libraries
- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse.
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- Interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl. Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0
You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries
- you have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- Interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0
## Dataset Model

### FT Document

1 Row for each:
- General_ID - Key
- OJV
- Source
- Place

Why? Because, for each OJV, we can detect a multi-place Vacancy  
 e.g. «Software Developer in London / Liverpool»

### FT Skill Analysis

1 Row for each:
- General_ID - Key
- OJV
- Source
- Place
- Skill

Why? Because, for each OJV, we can obviously detect multiple skills  
 e.g. «Software Developer in London / Liverpool, with customer orientation culture, that speaks english and tolerates stress»
Dataset Model

So, to have the number of job vacancies you always have to compute an unique count by General_ID
Tools

- Jupyter
- Apache Spark
- scikit-learn
- Hue
- Hive
- RStudio
- Amazon EMR
- Amazon Athena
- H2O.ai
- ETF (European Training Foundation)
AWS Athena Console

Data source
- AwsDataCatalog

Database
- etf_presentation_training

Tables and views
- Create

Filter tables and views

Tables (2)
- ft_document_en
- ft_skill_analysis_en
Live session with python and SQL