

UNIVERSITATEA TEHNICĂ „Gheorghe Asachi” din IAȘI
FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE
DOMENIUL: Calculatoare și tehnologia informației
SPECIALIZAREA: Tehnologia informației

Raport Intermediar I

Aplicație web pentru analiza CV-urilor

Profesor coordonator
Cristian Aflori

Studentă
Diana Secrieru

Iași, 2018

1. Scopul temei alese

În ziua de astăzi, domeniul resurselor umane este din ce în ce mai greu de gestionat, datorită unei creșteri a numărului de locuri de muncă și implicit a numărului de candidați. Tema aleasă urmărește dezvoltarea unei aplicații web ce are ca scop ușurarea procesului de recrutare, din perspectiva personalului de resurse umane. Această aplicație va permite analizarea CV-urilor și redarea conținutului acestora sub un format mai ușor de interpretat. Astfel, informațiile din sute de CV-uri vor fi transpuse sub forma unor diagrame, grafice, etc., ce ilustrează clasificarea candidaților în funcție de aptitudini, experiență anterioară de muncă etc. Scopul final este acela de a selecta acei candidați care se potrivesc cel mai bine cu specificațiile postului la care au aplicat.

2. Referințe la teme/subiecte similare

Printre temele/subiectele similare cu tema aleasă, se regăsesc următoarele:

- Managementul riscului - Indiferent de industrie, analiza insuficientă a riscului duce de cele mai multe ori la eșec. Acest lucru se aplică în mod deosebit în domeniul financiar, unde adoptarea de software pentru managementul riscului, bazat pe tehnologii de analiză de text, pot crește în mod dramatic abilitatea de a diminua riscul, prin oferirea abilității de a avea acces complet asupra informațiilor potrivite la momentul potrivit.
- Prevenirea criminalității informatice – Natura anonimă a internetului și a mediilor de comunicare ale acestuia contribuie la riscul crescut de crime informatice. Astăzi, aplicații ce folosesc tehnologii de text mining, fac prevenirea crimelor informatice mult mai ușoară pentru agențiile de informații.
- Serviciul de asistență pentru clienți – Analiza datelor, precum și procesarea limbajului natural, sunt aplicații frecvente în serviciul de asistență pentru clienți. Astăzi, software-ul de analiză a textului este deseori adoptat pentru a îmbunătăți experiența clienților, folosind diferite surse de informații valoroase, cum ar fi: chestionare, tichete de asistență, etc. Scopul este de a îmbunătăți calitatea, eficacitatea și viteza în rezolvarea problemelor clienților.

3. Resurse software utilizate

Pe partea de backend, aplicația va folosi tehnologii precum:

- Java
- Spring
- Hibernate
- Servlet

Pe partea de frontend, tehnologiile utilizate vor fi:

- HTML
- CSS
- JavaScript
- Bootstrap
- Angular
- jQuery
- AJAX
- JSON
- XML

Pentru stocarea datelor, va fi folosită o bază de date relațională : SQL.

Ca și medii de dezvoltare, vor fi utilizate: Eclipse (pentru dezvoltare Java), SQL Developer (client SQL).

4. Algoritmi sau metode alese

Dimensiunea datelor crește exponențial pe zi ce trece. Aproape toate tipurile de instituții, organizații și industrii de afaceri, stochează datele în format electronic. O cantitate uriașă de text trăiește pe internet, sub forma librăriilor digitale, depozitelor și a altor informații textuale, cum ar fi blog-urile, rețelele de socializare și email-urile. Este o sarcină provocatoare aceea de a determina tiparele potrivite și trendurile, pentru a extrage cunoștințe valoroase din acest volum mare de date. Uneltele tradiționale de analizare a datelor sunt incapabile de a se ocupa de datele textuale, deoarece este nevoie de timp și efort în extragerea informațiilor.

Metoda principală aleasă este analiza textului (text mining). Aceasta reprezintă un proces de extragere a tiparelor semnificative și interesante, pentru a explora cunoștințele ce reies din sursele de date textuale. Analiza textului este un domeniu multidisciplinar, bazat pe obținerea de informații, analiza datelor, învățarea automată, statistică și lingvistici computaționale. Printre tehnicile analizei de text se regăsesc: sumarizarea, clasificarea, clustering-ul, etc. Acestea pot fi aplicate pentru a extrage cunoștințe. Analiza textului se ocupă cu textul scris în limbaj natural, ce este stocat în format semistructurat și nestructurat. Tehnicile de analiză de date sunt aplicate în industrie, aplicații web, internet și alte arii. Domeniile de aplicație, precum motoarele de căutare, analizarea sugestiilor de produse, detectarea fraudelor și analizarea rețelelor de socializare, folosesc analiza datelor pentru exploatarea opiniilor, extragerea de caracteristici, analizarea predictivă etc.

Extragerea de informații valoroase în diferite documente este o sarcină plictisitoare și obositoare. Selecția tehnicilor potrivite pentru analiză de text reduc din timpul și efortul pentru a găsi tipare relevante de analiză și luare de decizii.

5. Rezultate așteptate

Aplicația va avea două interfețe, una pentru candidat, și una pentru personalul de resurse umane. Candidatul va avea posibilitatea de a-și crea un cont, iar în portalul aferent, acesta va putea să își încarce CV-ul. De asemenea, acesta va putea vizualiza statusul procesului de recrutare în care este implicat. Personalul de resurse umane se poate autentifica în portalul aferent, acesta oferind posibilitatea de a vizualiza atât statistici privind posturile disponibile cât și candidații ce au aplicat la fiecare job în parte. Angajații resurselor umane vor putea viziona sub un format mai ușor de interpretat, conținutul CV-urilor angajaților. De asemenea, există posibilitatea de a vedea care sunt cei mai potriviți candidați pentru un anumit post, în funcție de algoritmul de analiză a CV-urilor. Acesta presupune parsarea conținutului CV-urilor și extragerea cuvintelor cheie, care vor fi folosite la clasificarea candidaților și potrivirea aptitudinilor, experienței de muncă etc. cu specificațiile job-ului în cauză.