



Automatic Extraction of Nursing Tasks from Online Job Vacancies

Vladimer Kobayashi (v.kobayashi@uva.nl)

Dr Stefan Mol (S.T.Mol@uva.nl)

Dr Gábor Kismihók (g.kismihok@uva.nl)

Maria Hesterberg (Maria.Hesterberg@ukb.uni-bonn.de)



Overview

- Marching order
- Nursing tasks from vacancies
- Dealing with thousands of vacancies – Text Mining
- Text Mining: the nuts and bolts
- Our data
- Results
- Why challenging
- Limitations
- Future Work and... we need your help!

Marching order

- Extract nursing tasks from vacancies
- Is it even possible?

Pflegerische Leitung (m/w) für die Notaufnahme in Vollzeit

IHRE AUFGABEN: Personalführung und Einsatzplanung unter Berücksichtigung der gesetzlichen, tariflichen und betrieblichen Vorgaben | Fachliche Leitung und Sicherstellung der Pflegequalität | Koordinierung der Patientenversorgung sowie die Sicherstellung einer reibungslosen Zusammenarbeit | Kooperative Zusammenarbeit mit anderen Berufsgruppen und externen Dienstleistern | Aktive Beteiligung am Qualitätsmanagement | Vertrauensvolle Zusammenarbeit mit dem Pflegedirektorium

IHR PROFIL: Sie sind examinierter Gesundheits- und Krankenpfleger mit mehrjähriger Berufserfahrung und haben Erfahrung im Bereich der Notfallmedizin | Sie verfügen über eine Weite soziale Kompetenz und Lösungsansätze mit allen Beteiligten

WIR BIETEN: Fortbildungsmöglichkeiten, vom Therapeuten zum Krankheitspfleger
 Unsere Arbeitsplätze sind menschenorientiert



Landkreis
Krefeld • Münster
• Köln
• Bonn/Rhein
• Aachen

Wir leben unsere Werte



- Sie sind ein mutiger, aufrechter Verfechter von **Bewohnerinteressen?**
- Für Sie ist Alltagsgestaltung eine Kernaufgabe Ihrer täglichen Arbeit, neben Behandlungs- und Grundpflege?
- Sie sind ein streitbarer Geist, der nach modernen wissenschaftlichen Erkenntnissen arbeitet und gerne bereit ist, „alte, überholte Zöpfe abzuschneiden“?
- Sie schätzen gute Rahmenbedingungen wie eine leistungsgerechte Bezahlung, verlässliche Dienstplangestaltung und die Vereinbarkeit von Beruf und Familie?
- Sie bringen Erfahrungen im ambulanten Bereich mit, in der Führung von Teams und haben fundierte Kenntnisse im Bereich SGB V und XII?
- Dann sind Sie bei uns richtig – melden Sie sich gerne mit Ihren aussagekräftigen Bewerbungsunterlagen.

Gestalten Sie unsere Werte mit
 Senden Sie uns bitte Ihre aussagekräftigen Bewerbungsunterlagen unter Angabe Ihres frühestmöglichen Eintrittsdatums an jobs.waldhausen@gda.de oder per Post an unsere Adresse. Die Übernahme von Fahrtkosten und Verdienstaussfällen ist leider nicht möglich.



www.gda.de

GDA Wohnstift Hannover-Waldhausen
 GDA Service GmbH
 Stiftdirektor Benjamin Knollmann
 Hildesheimer Straße 183 · 30173 Hannover

11x in Deutschland



tasks

- Since Vacancies may contain job task information
So it is possible to extract job tasks.



- Extraction can be done manually

- Enough for hundreds of vacancies
- For thousands of vacancies, increase the number of manual coders.

➤ **But how about for millions of vacancies?**

- Still more manual coders? However reliability can become an issue. Cost of hiring? Cost of using resources? Time constraints?... **In other words manual coding is expensive and slow.**

Text Mining to the rescue

- Manually code a few hundred vacancies by sorting sentences (or phrases) into task or non-task categories.
- Use a text mining technique, specifically text classification, to train (or create) a classifier that can automatically classify sentences in a vacancy into task or non-task categories
- Ask an expert if the model is unsure about to which category a specific sentence should be assigned to.

Job vacancies posted online

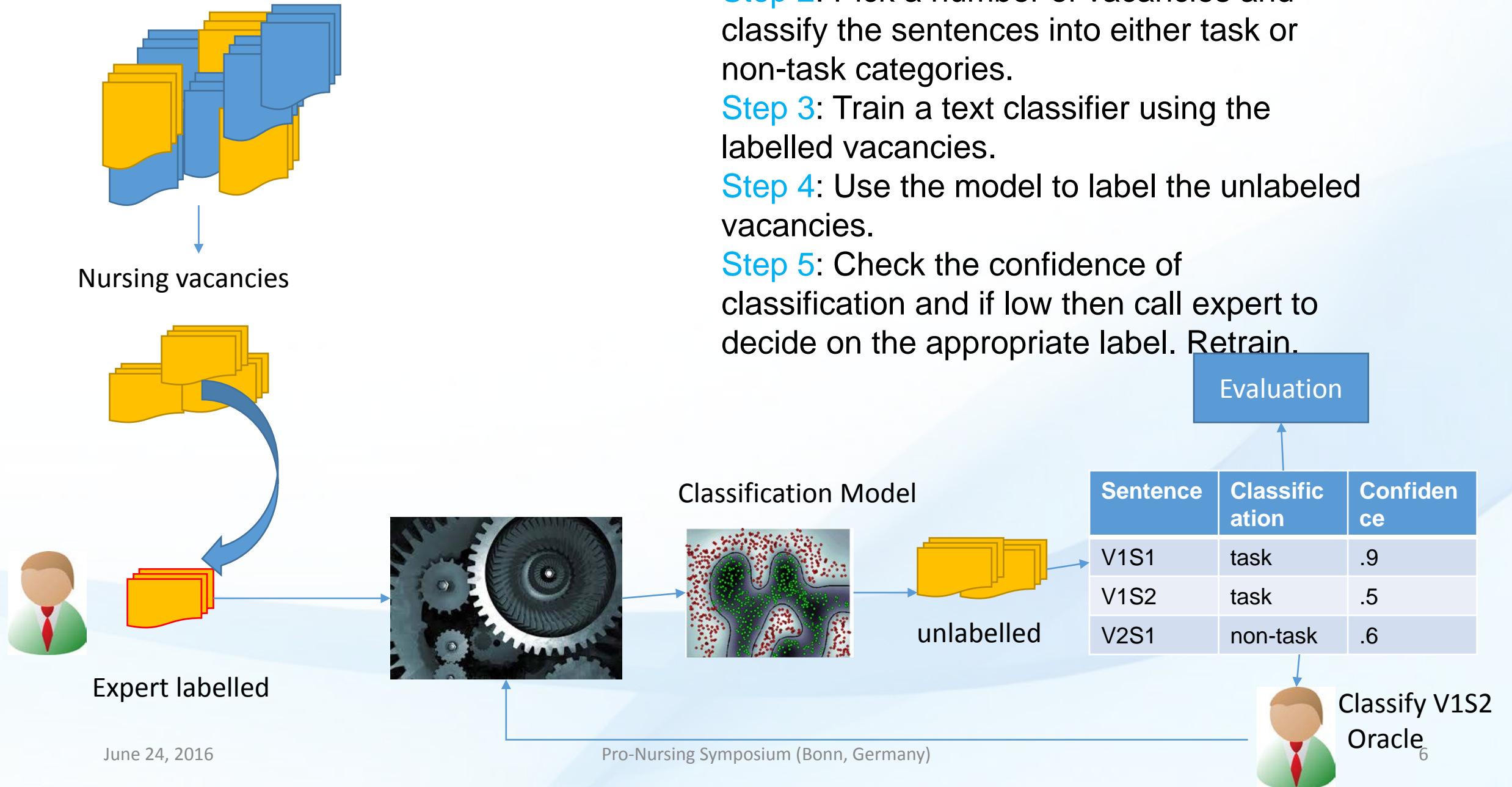
Step 1: Identify nursing vacancies.

Step 2: Pick a number of vacancies and classify the sentences into either task or non-task categories.

Step 3: Train a text classifier using the labelled vacancies.

Step 4: Use the model to label the unlabeled vacancies.

Step 5: Check the confidence of classification and if low then call expert to decide on the appropriate label. Retrain.



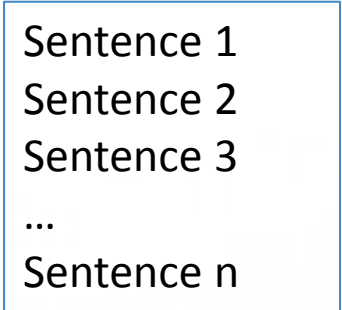


- Representation = feature + classifier
- Scoring
- Optimization

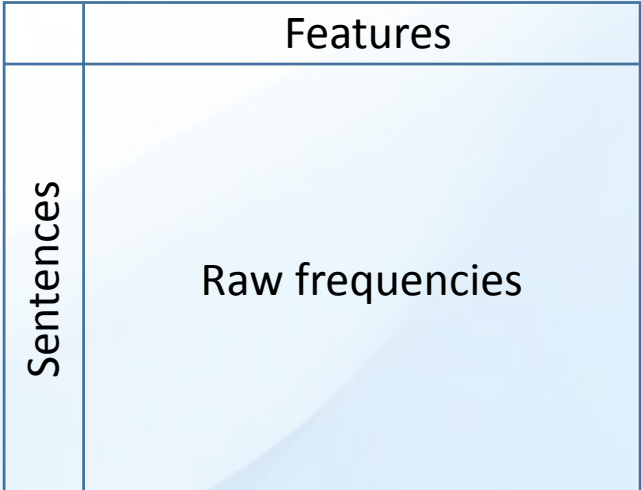
Consider a single vacancy



Preprocessing
Sentence Segmentation

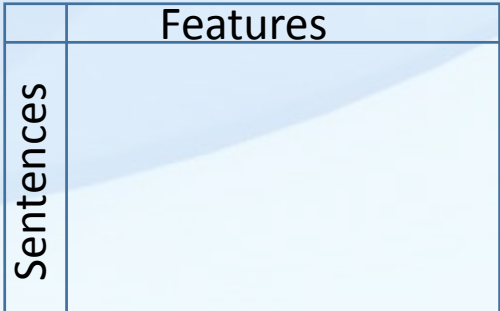


Transformation



Words are used as features

Dimensionality reduction

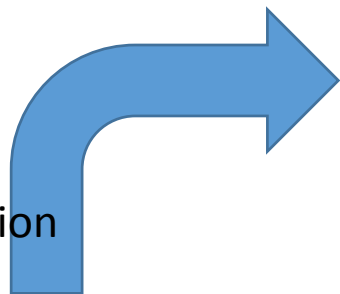


Final aggregated classification with confidence

Aggregate

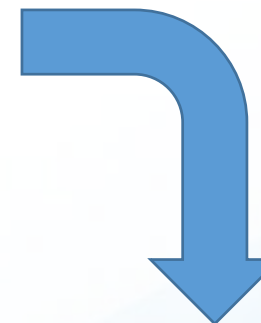
- Naïve Bayes
- Random Forest
- Support Vector Machines

Preprocessing
Sentence Segmentation



Die Alexianer sind ein Unternehmen der Gesundheits- und Sozialwirtschaft, in dem bundesweit rund 12. Die Alexianer St. Hedwig Kliniken Berlin ist ein katholischer Krankenhausträger und Teil der Un In unserer Notaufnahme versorgen wir ca. 25.000 Patienten pro Jahr mit chirurgischen, internistische
Pflegerische Leitung (m/w)
für die Notaufnahme
in Vollzeit
IHRE AUFGABEN: Personalführung und Einsatzplanung unter Berücksichtigung der gesetzlichen, tariflich
IHR PROFIL: Sie sind examiniertes Gesundheits- und Krankenpfleger mit mehrjähriger Berufserfahrung u
WIR BIETEN:
Ein interessantes, entwicklungsfähiges Aufgabengebiet | In- und externe Fortbildungsmöglichkeiten |
Unsere Arbeitsatmosphäre ist geprägt durch ein Miteinander, das durch das christliche Menschenbild b
Für Rückfragen steht Ihnen unsere Pflegedirektorin, Frau Reimer, unter 030/6741-2500 gerne zur Verfö
Bitte richten Sie Ihre vollständige Bewerbung (Anschreiben, Lebenslauf, Zeugnisse) unter Angabe der
Alexianer Service GmbH
Personalabteilung
Zentrales Bewerbermanagement (ZBM)
Große Hamburger Straße 3
10115 Berlin
Telefon: 030/400372-604
Fax: 030/400372-599
Weitere Informationen zu den Einrichtungen des Unter-nehmens erhalten Sie im Internet unter:
www.alexianer.de

Transformation



Pflegerische Leitung (m/w) für die Notaufnahme in Vollzeit

IHRE AUFGABEN: Personalführung und Einsatzplanung unter Berücksichtigung der gesetzlichen, tariflichen und betrieblichen Vorgaben | Fachliche Leitung und Sicherstellung der Pflegequalität | Koordinierung der Patientenversorgung sowie die Sicherstellung einer reibungslosen Zusammenarbeit | Kooperative Zusammenarbeit mit anderen Berufsgruppen und externen Dienstleistern | Aktive Beteiligung am Qualitätsmanagement | Vertrauensvolle Zusammenarbeit mit dem Pflegedirektorium

IHR PROFIL: Sie sind examinierter Gesundheits- und Krankenpfleger mit mehrjähriger Berufserfahrung und haben Erfahrung im Bereich der Notfallmedizin | Sie verfügen über eine Weiterbildung für Leitungsfunktionen im Gesundheitswesen | Sie haben hohe soziale Kompetenz in Verbindung mit der Fähigkeit, Mitarbeiter zu kreativen Lösungsansätzen im Team zu motivieren | Sie sind zu berufsübergreifender Kooperation mit allen Mitarbeitern unseres Hauses bereit | Sie verfügen über gute EDV-Kenntnisse

WIR BIETEN: Ein interessantes, entwicklungsfähiges Aufgabengebiet | In- und externe Fortbildungsmöglichkeiten | Kostenlosen Zugang für CNE (Certified Nursing Education) vom Thieme Verlag | Ein angenehmes Arbeitsklima | Vergütung nach AVR/C | Eine leistungsstarke betriebliche Altersvorsorge | Arbeitgeberfinanzierte Betreuung bei Krankheit des Kindes

Unsere Arbeitsatmosphäre ist geprägt durch ein Miteinander, das durch das christliche Menschenbild bestimmt ist. Wir wünschen uns dabei Ihre aktive Mitwirkung.

Für Rückfragen steht Ihnen unsere Pflegedirektorin, Frau Reimer, unter 030/6741-2500 gerne zur Verfügung.

Bitte richten Sie Ihre vollständige Bewerbung (Anschreiben, Lebenslauf, Zeugnisse) unter Angabe der Referenznummer BHHH14OPGDRONA2310 an perspektiven@alexianer.de oder an:

Alexianer Service GmbH
Personalabteilung
Zentrales Bewerbermanagement (ZBM)
Große Hamburger Straße 3
10115 Berlin
Telefon: 030/400372-604
Fax: 030/400372-599

Weitere Informationen zu den Einrichtungen des Unternehmens erhalten Sie im Internet unter: www.alexianer.de

Terms	Docs	beteiligung	betreutes	betriebliche	betriebsklima	betriebsvereinbarung
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

Terms	Docs	bewerben	bewerberinnen	bewerbermanagement	bewerbung
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0

Terms	Docs	bewerbungenasmittelhessende	bewohnerinteressen	bewohnerinnen	bezahlung
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	1	0	0
6	0	0	0	0	0

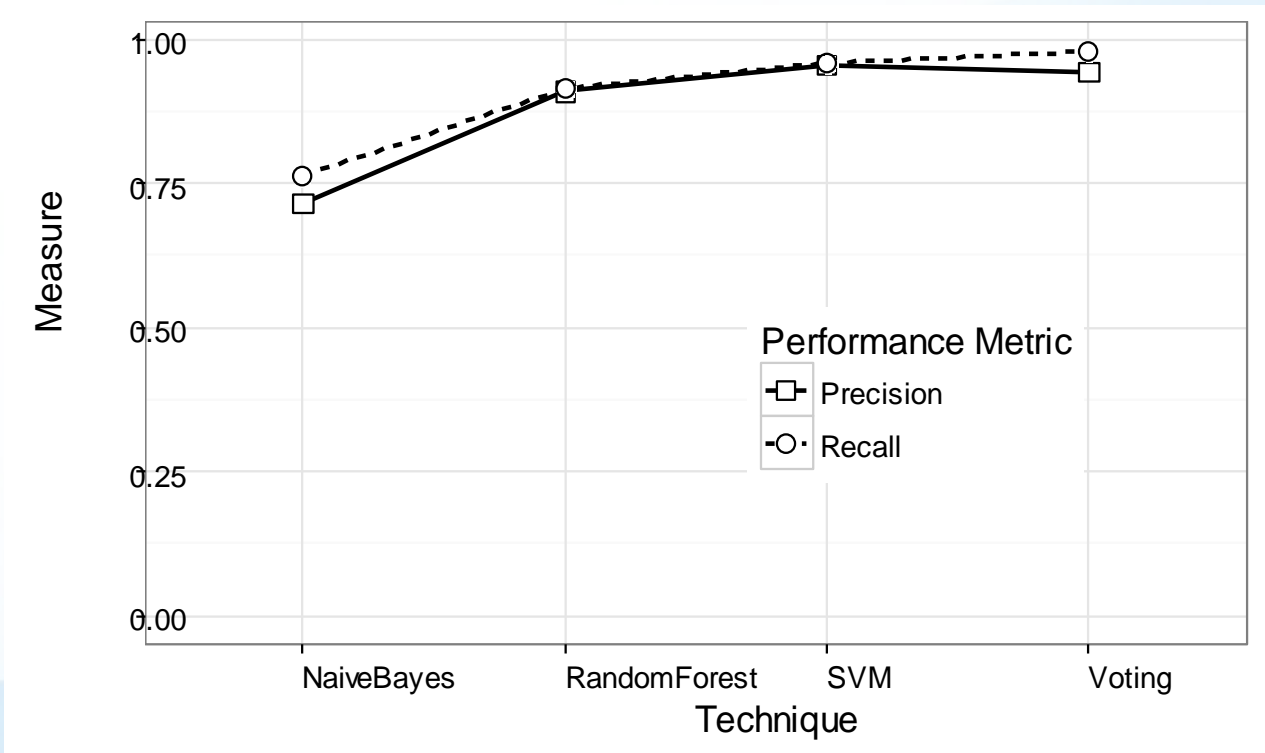
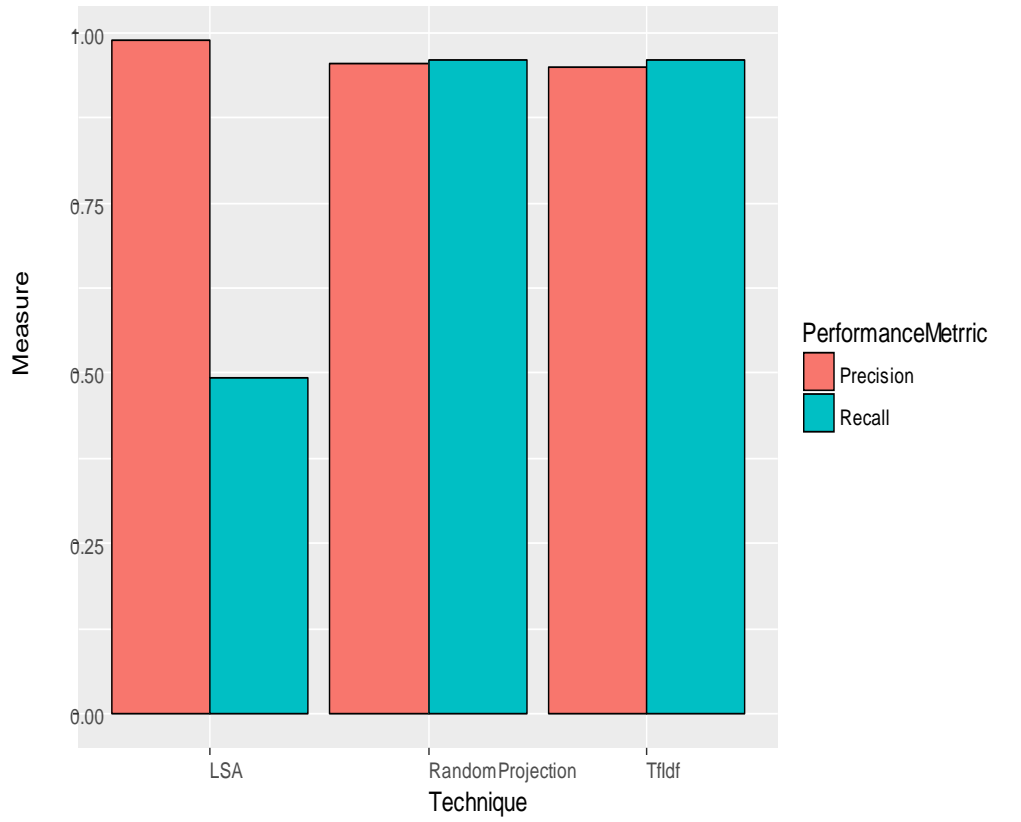
Terms	Docs	bezeichnungen	beziehen	bieten	bildau	bildquelle	bitte	boyenstr	bringen
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0

Our data

- 14000 vacancies (from Monster)
- 6000 nursing vacancies (about 40%), 180000 sentences
- Initial training data:
 - approx. 2000 labelled sentences
 - Feature size is 13000 (vocabulary)
 - Only 7% are task sentences (i.e. less than 150 sentences)
- We obtained additional 2000 task sentences
- Model performance: highest precision is .80
- Since there are many redundant task sentences we clustered the sentences and obtained 90 clusters

Results

Trained using SVM



Dimensionality Reduction

Model Performance

ID#	TASK	TASK CLUSTER	GERMAN TRANSLATION
	<i>Activity performed by nurse</i>	<i>Groupings of tasks that go together</i>	<i>Activity performed by nurse</i>
1	Improving interdisciplinary cooperation and collaboration	TEAMWORK	Verbesserung von interdisziplinärer Zusammenarbeit
2	Communicating with patients, family, and staff	TEAMWORK	Kommunikation mit Patienten, Angehörigen und Kollegen
3	Monitoring the patients' therapy	BASIC CARE	Überwachung der Therapie des Patienten
4	Assisting and monitoring anesthesia	SPECIAL CARE	Assistenz und Überwachung der Narkose
5	Assisting at surgical interventions	MEDICAL CARE	Assistenz bei operativen Eingriffen
6	Caring for patients with windpipe	SPECIAL CARE	Pflege von Patienten mit Tracheostoma
7	Checking vital signs	SPECIAL CARE	Vitalzeichenkontrolle
8	Administrating of medication	SPECIAL CARE	Verabreichung von Medikation
9	Providing nursing documentation	INTERNAL MANAGEMENT	Durchführung der Pflegedokumentation
11	Giving patients and family advise	PATIENT MANAGEMENT	Beratung von Patienten und Angehörigen
12	Participating in meetings	TEAMWORK	Teilnahme an Arbeitssitzungen
13	Working on doctors' prescriptions	PATIENT MANAGEMENT	Umsetzung von ärztlichen Anordnungen
14	Realising the nursing process	PATIENT MANAGEMENT	Umsetzung des Pflegeprozesses
15	Developing an individual care planning	PATIENT MANAGEMENT	Entwicklung eines individuellen Pflegeplans
16	Realising the wound management	SPECIAL CARE	Durchführung des Wundmanagements
17	Participating in hygiene training	EDUCATION	Beteiligung an Hygienefortbildungen
18	Doing laboratory tests	MEDICAL CARE	Durchführung von Labortests
20	Auditing medical transcriptions	INTERNAL MANAGEMENT	Prüfung und Umsetzung von medizinischen Anordnungen
21	Organizing the work on the ward	INTERNAL MANAGEMENT	Organisation der Stationsarbeit
22	Caring for the elderly	BASIC CARE	Pflege von älteren Menschen
23	Processing instruments	INTERNAL MANAGEMENT	Aufbereitung von Instrumenten
24	Condensing the latest results of nursing research	EDUCATION	Miteinbeziehen des aktuellen Standes der Pflegewissenschaft
25	Providing basic or general care	BASIC CARE	Durchführung der Allgemeinen Pflege

Why Text Mining works

- Word and word combinations provide a sufficient representation for this task
- We can even identify *task words* (e.g. aufgaben, pflege)
- Leverage big data and big data tools

Why this is challenging?

- It is costly to obtain labelled data
- More non-task compared to task sentences (unbalanced data)
- Needs validation: are the extracted tasks useful from the point of view of job holders, experts and for other applications (e.g. job analysis)

Limitations of our approach

- Does not take into account the order of sentences in a vacancy.
- We need more data (specifically labelled data).
- The model is specific to the German language. Though the tasks can be translated to English or we can easily adapt the analysis pipeline to other languages.

Future Work and Possible help

- Collect additional data. **Do you have data to share?**
- Help from experts. **Are you familiar with the nursing profession? Help us in the labeling or be an oracle and participate in our surveys.**
- Try other analytical methods such as sequence labeling.
- Fine tune parameters using other evaluation metrics.
- Try semi-supervised learning
- Improved the active learning approach
- Choose a different representation that does not rely too much on labelled data
- Crowd-sourcing

Key takeaways

- Text mining works given an abundance of data, good representation and quality of labelled data.
- Expert plays an important role (e.g. active learning)
- Random Projection and tf-idf scoring work well for dimensionality reduction
- Combining classifiers is better than just adopting a specific classification technique.
- Validation is important for text mining results.

Acknowledgement

- Pro-nursing project (<http://www.pro-nursing.eu/web/>)
- Monster (<http://www.monster.com/>)
- Eduworks Network (<http://www.eduworks-network.eu/>)
- European Commission
- For questions and support kindly email me at v.kobayashi@uva.nl or email Marjan Khobreh of the Pro-nursing project