

Predicting Employee Expertise for Talent Management in the Enterprise

Kush R. Varshney, Vijil Chenthamarakshan, Scott W. Fancher,
Jun Wang, Dongping Fang, Aleksandra Mojsilovic

KDD2014読み会

京都大学石井研究室 B4

児玉 悠

概要

- IBMのIndustry paper
- 変化の激しい大企業内での人材マネジメント促進のためには最新の社員の専門領域に関するデータが必要→データの生成に機械学習を用いて自動化する
- webリソースや企業内データに対して機械学習を適用し社員の専門領域クラスラベルを予測する

背景

- IBM: 情報技術やそれに基づくビジネスコンサルを行う
巨大企業←全世界の社員合計は425000人
- 情報系企業の技術革新のペースは速くなる一方
- IBMのような巨大企業が変化の激しいトレンドを反映した意思決定を行うためには、正確かつ最新の各社員の専門性に関する情報が必要
→機械学習を用いて情報の整理を自動化！

機械学習での定式化

- 社員の専門領域クラスラベルの推定を教師有り分類問題として捉える
- ここではSales部門の社員約36000人を11クラスの専門領域に分類することを考える
- 使用するデータ・特徴量・アルゴリズムはどうか？

使用できるデータ

- 使用可能と思われるデータソースは次の4種類
 - 社員自身の仕事の説明文データ(job titleデータ)
 - 標準的な人事情報(HRデータ)
 - 各社員のプロダクトの情報→職種に依存するため用いない
 - IBMの社内SNSからのデータ

Tags	
Industries	
2 government	+1
2 healthcare	+1
2 retail	+1
Clients	
Products and Services	
Other Tags	
3 expertise-locator	+1
3 research	+1
3 smarter-workforce	+1
2 advanced_analytics	+1
2 analytics	+1
View All (38)	

(a)

Opportunity Name and Description	Opportunity Date
SW & Systems Assessment/Roadmap	2-Sep-13
MOI- Portal with e-services	10-Feb-13
P series for Research	2-Dec-13
Leadership Development - Ashghal	19-May-13
Dynamic Infra Assessment	15-Sep-13
Programme Management Assistance	18-Jun-13
Programme Management Assistance	18-Jun-13
Security Framework & Next Generation SOC	2-Sep-13
ESB- QF Wide project	20-Nov-13
MOI- Portal with e-services	10-Feb-13
Mobile Solutions- HR	5-Nov-13
Guardium	20-Nov-13
QF- Portal	10-Feb-13
Predictive Maintenance and Quality for QatarGas - As pare of the enterprise	30-Sep-13
Predictive Maintenance and Quality Extending the Asset Management implement	30-Sep-13
Predictive Maintenance and Quality	27-Nov-13
I2 Oppt. Client is working on building a scope of work for managing internal	13-May-13
Security Advisory Services GTS	14-Nov-13
Predictive Asset Maintinace and Quality. Part of the Enterprise Architectur	6-Oct-13
TOM and Busines Process Services	8-Jul-13
SAP Consolidation SAP PS module	8-Jul-13
Predictive Analytics - PMQ(New) As part of the DW/BI project	27-Nov-13
New Platform for Scoring system in Qatar	8-Oct-13
QAPCO SAP HW Platform	22-Sep-13
Business Process Management Consulting Services	15-Sep-13

(b)

John Smith

Job Title

Workforce Diversity & Programs Professional: Workforce Analytics & Initiatives

What I'm Known For

"Deep experience in Enterprise Workforce Management, Expertise Data and Reporting, EA Tool, IBM Expertise, and Sales Eminence Initiative and Sales Eminence Top Sheet management system. Also lead the Sales Expertise Management Initiative (SEMI) community"

(c)

Figure 1: Examples of raw data for a single employee: (a) tags from IBM Connections, (b) sales opportunity data, and (c) job title.

特徴抽出

- 使用するの基本的にはテキストデータである(job title データ、HRデータ、sales opportunity データ、社内 SNSのソーシャルタグデータ)
- 用いる特徴量は上記データからのbag-of-words

使用するアルゴリズム

- 以下4つのone-vs-restを分類器として精度を比較
 - L2正則化ロジスティック回帰
 - L1正則化ロジスティック回帰
 - SVM
 - ナイーブベイズ

実験結果1

□ 各アルゴリズムの比較

□ 各特徴セットの比較

→Job Title + HR Info にロジスティック回帰を適用した場合の精度が最も良い

Table 1: Fivefold Cross-Validation Test Accuracy

Feature Set	ℓ_2 -Reg. Logistic Regress.	ℓ_1 -Reg. Logistic Regress.	Support Vec. Mach.	Naïve Bayes
Job Title (A)	0.6746	0.6749	0.6695	0.6410
HR Info (B)	0.7661	0.7641	0.7604	0.6807
Social Tags (C)	0.2320	0.2396	0.2380	0.2573
Sales Opp (D)	0.3374	0.3404	0.3473	0.2306
(A) + (B)	0.8016	0.8031	0.7899	0.7330
(A) + (B) + (C)	0.7671	0.7703	0.7504	0.6118
(A) + (B) + (C) + (D)	0.7720	0.7733	0.7655	0.3952

実験結果2

- 適合度が高い k 個の推定クラスの中に正解クラスが入っている確率

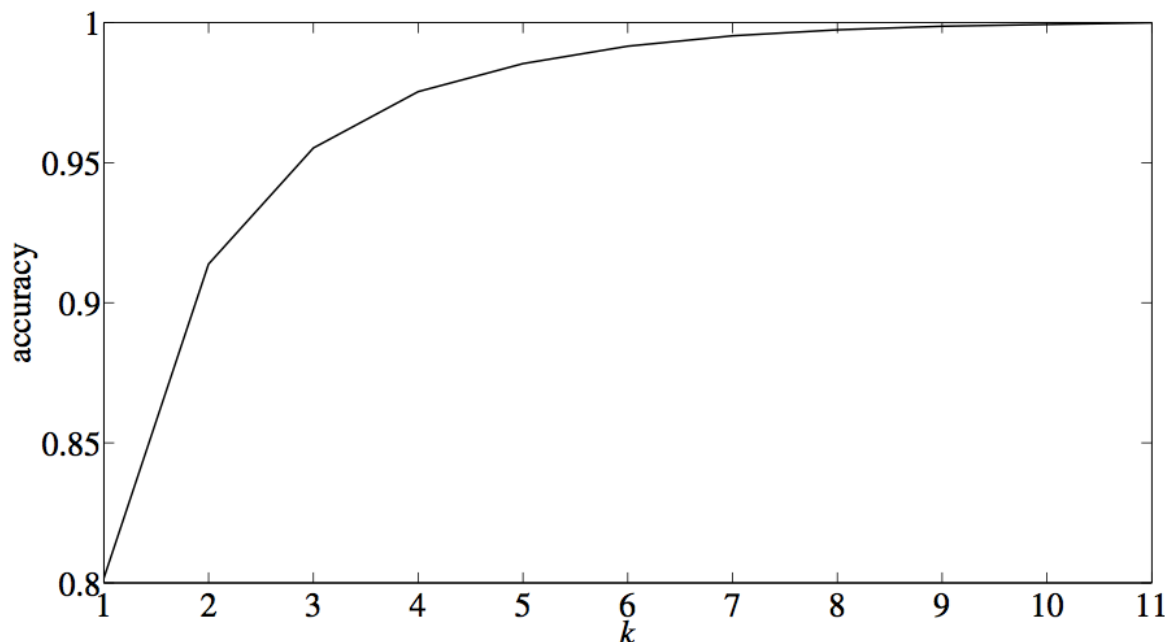


Figure 2: Fivefold cross-validation accuracy as a function of number of predictions per employee.

展開と影響

- IBMは実際に今回の実験より得られた結果を用いてIBM社員の専門領域推定システムを構築しようとしている
- その開発と普及はまだ途上段階であるが、実現すれば手作業による分類のコストの大幅な削減が期待できる

実際に開発したシステム画面

The screenshot displays the IBM Workplace Job Selector interface. At the top, there's a navigation bar with 'w3' logo, 'IBM Workplace', and various links like '1-3-9', 'My Links', 'You & IBM', 'Collaboration', and 'Support'. A search bar on the right says 'w3 Search Pages(w3)'. Below this is a blue header with 'Job Selector' and a user greeting 'Welcome back, Takumi Hatomi'.

The main content area has tabs for 'Manager Dashboard', 'Employees Requiring Action (3)', and 'All Employees (10)'. A message states: 'Below are a list of your employees and their current job and job recommendations.' There are buttons for 'Confirm job' and 'Change job...'.

The table below lists three employees with their status, current job role specialty, and recommendations.

☐	Name	Status	Current Primary Job Role Specialty	Actions
☐	Adams, Emily	Manager Action Required - Job change recommended - Last updated: 20 Jan 2013 Flagged for Skill Focal Assistance	12% Customer Service Representative: Client Service Center ▶ Recommendations (Expand to show 3 of 8. View All)	Actions...
☐	Branson, Zoe	Manager Action Required - Review required - Last confirmed: 16 Feb 2013	32% Customer Service Representative: Service Desk Delivery ▼ Recommendations (Showing 3 of 12. View All) 43% Customer Service Representative: Client Service Center 12% IMAC Coordinator:Service Desk Delivery 9% Technical Support Representative: Service Desk Delivery	Actions...
☐	Beck, Adam	Manager Action Required - Not aligned to short list - Last confirmed: 22 Mar 2013	9% Software Developer: COBOL ▼ Recommendations (Showing 3 of 17. View All) 38% Application Developer: Java 22% Application Developer: COBOL 15% Software Developer: Java	Actions...

At the bottom, it shows '1 - 3 of 3 items', '0 items selected', and pagination controls (1, 10, 25, 50, 100, All).

Figure 3: Wireframe of job role specialty prediction deployment.

まとめ

- 人材マネジメントのためのデータ整理を自動化するために、IBMの企業データに機械学習を適用して社員の専門領域クラスを推定した
- job titleデータとHRデータからのbag-of-wordsにロジスティック回帰を用いた場合が最も精度が良かった
- IBMでは実際にこのシステムを開発し普及させることでコスト削減&トレンドに即した意思決定の支援を行おうとしている