# Predicting Employee Expertise for Talent Management in the Enterprise

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





#### **Human Resource Analytics**



image source: Bersin (2013)



### **Talent Analytics**

- Largest worldwide employers today are knowledge-based enterprises
- Most important asset is human capital (Schultz, 196)
- Knowledge workers are unique, each have in aividuari expertise
- Most basic of problems: inventorying employees according



### Why a Talent Inventory is Needed

- Quickening pace of technological innovation
  - New products, solutions, and acquisitions emerge each quarter
- Important for strategic and tactical business decision-making to be informed by complete, precise, accurate, and up-to-date information on the expertise of employees (Hu, Ray, and Singh, 2007)
- Tactical example: What team should serve a given client (in terms of composition of employee skills)
- Strategic example: Which emerging technology areas does the company have the talent to support



### **Problem Statement**

 Develop predictive analytics based upon employees' digital footprints to constantly update the current inventory of expertise across an organization in a way that commingles with existing business processes



# **Relationship to Prior Work**

- LinkedIn and other similar skill recommendation systems have free-form skill description
  - Cannot integrate with existing ecosystems of processes and reporting tools built around expertise taxonomies
- Prior work on expertise prediction within an enterprise has been based only on internal social media data and has not integrated with business processes (Shami et al., 2009; Guy et al., 2013)

#### IBM

## **IBM Corporation**

- Focus on a deployed system within the IBM Corporation
- Approximately 425,000 employees worldwide
  - Hardware, software, consulting services, research, sales, support, ...
- Five-level expertise taxonomy
  - Sample values on next slide
- Employees assess themselves against the taxonomy
  - A significant fraction have incomplete, incorrect, or out-of-date assessments



# **IBM Expertise Taxonomy**

Taxonomy Level	Sample 1	Sample 2	Sample 3	
Primary Job Category	Sales	Human Resources	Research	
Secondary Job Category	Industry Sales	Learning	Research Staff	
Job Role	Brand Client Representative	Learning Consultant	Research Scientist	
Job Role Specialty	Brand Client Representative: BAO-Advanced Analytics & Optimization	Learning Consultant: Collaboration, Knowledge & Communities	Research Scientist: Computational Biology	
Skill	Sell ILOG Optimization	Analyze Performance Improvement Needs	Develop Algorithms for Biological Data Analysis	

Variables to Predict



# **Machine Learning Formulation**

- Treat job role or specialty prediction for an employee as a supervised classification problem
  - Very large number of classes
- Varied features derived from employees' digital footprints within the company
- Present top *k* predictions as output along with confidence value
- Misclassification error is in fact the most appropriate performance metric
- Experiment with one-versus-all  $\ell_1$ -regularized logistic regression,  $\ell_2$ -regularized logistic regression, SVM, naïve Bayes



### **Features**

- A. Employee-entered free text on their responsibilities
- B. Basic HR information
- C. Internal social media (tags, blogs, wikis, etc.)
- D. Job-specific data sources like sales opportunities for salespeople, publications for researchers, etc.



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"

2		D	
	=	Opportunity Name and Description	Opportunity Date
ags =		SW & Systems Assessment/Roadmap U-CO	2-Sep-13
ndustries		MOI- Portal with e-services	10-Feb-13
		P series for Research	2-Dec-13
government		Leadership Development - Ashghal	19-May-13
		Dynamic Infra Assessment	15-Sep-13
healthcare	-1	Programme Management Assistance	18-Jun-13
		Programme Management Assistance	18-Jun-13
retail	+1	Security Framework & Next Generation SOC	2-Sep-13
		ESB- QF Wide project	20-Nov-13
lients		MOI- Portal with e-services	10-Feb-13
		Mobile Solutions- HR	5 Nov-13
Products and Services		Guardium	20-Nov-13
		QF- Portal D	10-Feb-13
)ther Tags		Predictive Maintinance and Quality for QatarGas - As pare of the enterprise	30-Sep-13
		Predictive Maintenance and Quality Extending the Asset Management implement	30-Sep-13
expertise-locator +1		Predictive Maintinance and Quality	27-Nov-13
		I2 Oppt. Client is working on building a scope of work for managing internal	13-May-13
research	+1	Security Advisory Services GTS	14-Nov-13
		Predictive Asset Maintinance and Quality. Part of the Enterprise Architectur	6-Oct-13
smarter-workforce	+1	TOM and Busines Process Services	8-Jul-13
		SAP Consolidation SAP PS module	8-Jul+13
advanced_analytics +		Predictive Analytics - PMQ(New) As part of the DW/BI project	27-Nov-13
		New Platform for Scoring system in Qatar	8-Oct-13
analytics	+1	QAPCO SAP HW Platform	22-Sep-13
View /	AII (38)	Business Process Management Consulting Services	15-Sep-13



# **Empirical Study**

- Predicting job roles of salespeople
- 11 class problem (imbalanced)
  - Brand Client Representative (BCR), Client Representative (CR), Client Technical Architect (CTA), Client Technical Manager (CTM), Client Technical Specialist (CTS), Client Unit Executive (CUE), Industry Solution Representative (ISR), Mid-Market Client Representative (MCR), Solution Representative (SR), Solution Representative - Brand Specialist (SRB), and Solution Sales Manager (SSM)
- Approximately 37,000 employees in training set; 5,000 in test set



#### **Fivefold Cross-Validation Accuracy**

Feature Set	$\ell_2$ -Reg. Logistic Regress.	$\ell_1$ -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



#### **Cross-Validation Accuracy (** $\ell_2$ -Regularized Logistic Regression)



© 2014 International Business Machines Corporation

# **Deployment**







© 2014 International Business Machines Corporation

#### IBM

#### Impact

- Initial deployment (without interface) to obtain correct job roles for approximately 4,000 worldwide salespeople
- To get salespeople to correctly enter their expertise would have taken approximately 30 minutes per employee
  - Save 1 person-year of effort = \$1M of revenue generated by salesperson
- Estimate 20 person-years of effort savings when deployed to entire company for annual assessments
- No need to limit to annual assessment
  - Just-in-time inventories, point-of-sale inventory updates, economic order quantities, and predictive inventory demand become possible for human resources and expertise management

#### IBM

# Summary

- Talent and human capital is a knowledge-based company's most valuable resource that must be harnessed properly using trusted expertise information
- Developed a classification methodology to predict the expertise of employees based on features derived from the digital footprints of employees
  - Label set from expertise taxonomy
- In the process of deploying the system for use by IBM Corporation
  - Should result in approximately twenty person-years of savings in annual updates of job roles and specialties
  - Impact is even greater than the savings in manual effort, because all business processes that depend on complete, accurate, and updated expertise data benefit from the predictions
  - Because of the steep reduction in effort, it will now be possible to update expertise assessments much more frequently than once a year, which is a transformation required to compete in today's dynamic business environment

#### Questions



