

(A Few) Key Lessons Learned Building Recommender Systems For Large-scale Social Networks



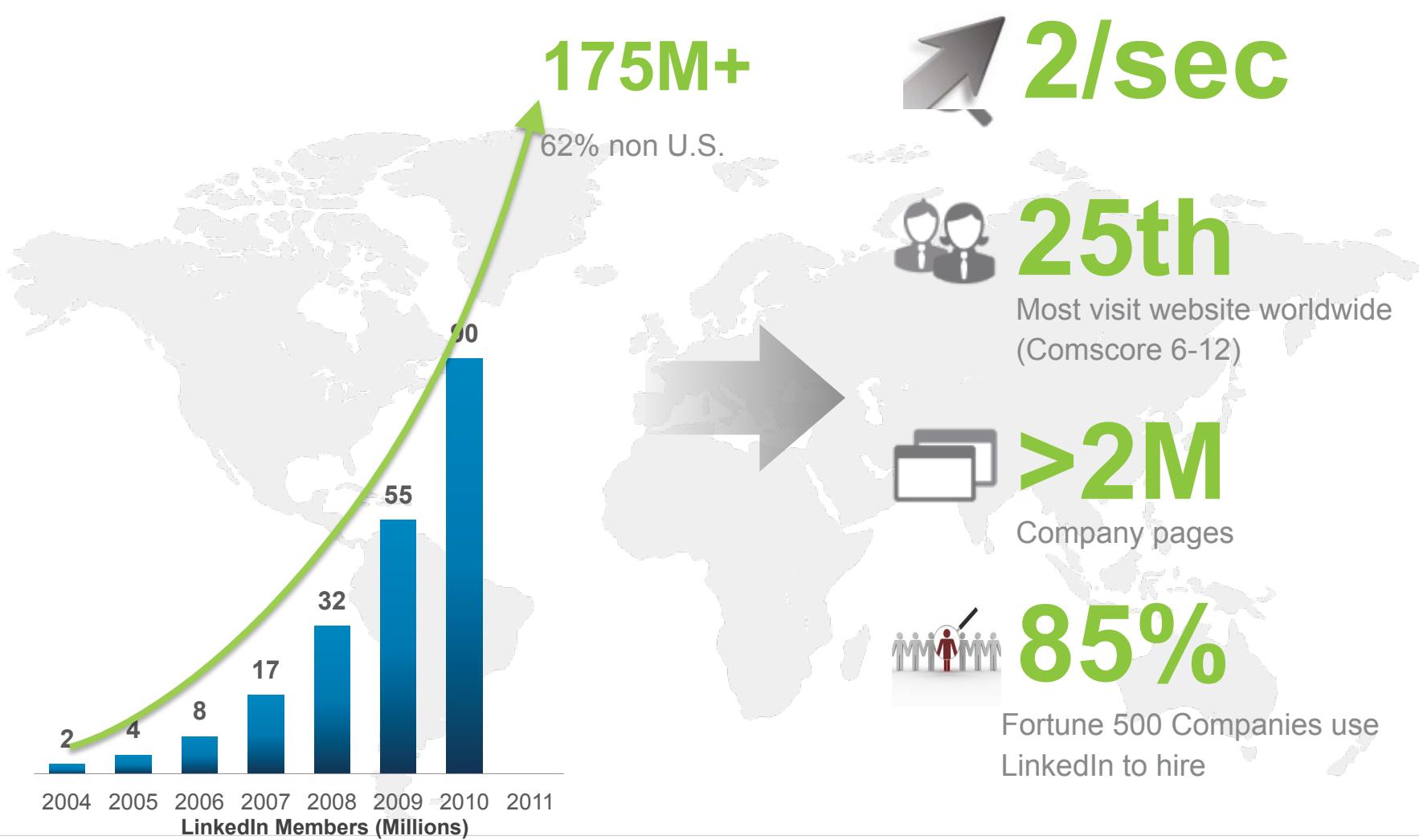
Christian Posse

Principal Data Scientist and Product Manager at
LinkedIn

San Francisco Bay Area | Information Technology and Services



World's Largest Professional Network



LinkedIn Homepage

The LinkedIn homepage features several key sections:

- Top Navigation:** Home, Profile, Contacts, Groups, Jobs, Inbox, Companies, News, More, People, Search, Advanced.
- Hero Section:** MIT Executive Education - MIT's Unique Entrepreneurial Education Program. Learn More & Enroll Today.
- Share an update:** A placeholder for users to share content.
- LinkedIn Today:** See all Top Headlines for you. Examples include "Wikipedia Blackout: Survive with These 12 Alternatives", "'No' is the New 'Yes': Four Practices to Reprioritize Your", and "Yahoo Co-Founder Jerry Yang Resigns".
- Recent Updates:** A feed of recent connections and activity. Examples:
 - Chad Whitney is now connected to Benjamin Menasha, Consulting Engineer at ITA Software and Earl-Bert Parco, Telecom Technician at LinkedIn
 - Jay Kreps is now connected to Jack Hutchins, Tugboat Captain at Westar Marine Services and Jack Wu, at LinkedIn
 - Peter Skomoroch is now connected to Aaron Cordova, Software Engineer at The Interactive, Inc. and Monica Nguon, Senior Regulatory Consultant at Deloitte & Touche
 - Peter Skomoroch Ignite Video: @mattwate Explains the Product Management Lessons of Battlestar Galactica <http://lnkd.in/MKd-hn>
 - Management Wisdom of Battlestar Galactica - Matt Wate youtu.be/mattwate Matt Wate presents, Management Wisdom of Battlestar Galactica, at Ignite Foo...
 - Kim Pope Shopzilis is looking for: Front End Engineer Learn more or Jobvite a friend <http://lnkd.in/2Bz5U>
 - Kay Luo via Twitter [http://tiny.cc/meyarw](#) Kuo Important Insights: Paper Tigers: What happens to all the Asian-American overachievers when the test-taking ends? <http://tiny.cc/meyarw>
 - Paul Whitney is now connected to Richard May, Director National Visualization and Analytics Center at Pacific Northwest National Laboratory and Alexandre Tantakovsky, scientist at Pacific Northwest National Laboratory
 - Andrew Pandre, Ph.D. is now connected to Jonathan Morales , Director, Decision Support
 - Bryan Machado is now connected to Sean Patten, IT Specialist, Devender Babu, Vice President Hadoop Product Management and Worldwide Marketing at Alavat and Janelle Delgado, Technical Recruiter
- People You May Know:** A list of recommended contacts including Bhushan Heda, Shomit Ghose, Ajay Mathew, and Christian Posse.
- Get over 30 premium benefits and powerful spend capacity.** A call-to-action button labeled "LEARN MORE" and "OPEN".
- Who's Viewed Your Profile?** A chart showing 515 Connections, 100+ Profile Viewers, and 2 New Viewers since 01/08.
- Jobs You May Be Interested In:** A list of job recommendations including Bioinformatics/Senior XDs, Senior Staff Systems Engineer (Data), and Sr. Applied Scientist.
- Groups You May Like:** AnalyticBridge, Bayesian Belief Networks with BayesiaLab, and Sentiment Analysis Symposium.
- Companies You May Want To Follow:** ness, Linguamatics, STATISTICS & ANALYTICS CONSULTANTS, ontotext, and digital trowel.

Powered by
Recommendations!

The Recommendations Opportunity

Jobs You May Be Interested In

Talent Match

CAP

Similar Profiles

Companies

Recommendations, similar companies search, peer companies, and company browse maps, company products and services browse maps

Groups

Recommendations, similar groups search

Related search

Network updates

Profile browse maps

Connections

Similar jobs

Jobs browse maps

Ad matching engine

$pCTR = f(member, creative, advertiser, context, inventory, OCTR)$

Referral Engine

Events You May Be Interested In

News

LinkedIn Today: See all Top Headlines for You

Meme Machine: SOPA Edition

"No" is the New "Yes": Four Practices to Reprioritize Your

Yahoo Co-Founder Jerry Yang Resigns



What are they worth? Think 50%

- > 50% of connections are from recommendations (PYMK)
- > 50% of job applications are from recommendations (JYMBII)
- > 50% of group joins are from recommendations (GYML)

People You May Know updated

 **Sophie Perisic**, French Teacher at Crystal Springs Uplands School [Connect](#) X

 **Lissa Juan**, Senior Manager, Product Management - Mobile & [Connect](#) X

 **Jeff King**, Sr. Director X.commerce Platform Partnerships at eBay [Connect](#) X

[See more »](#)

Jobs You May Be Interested In beta

 **Director of Engineering** Pandora - Oakland, CA X

 **Director of Engineering for Core...** CyberCoders - Santa Clara, CA X

 **Director of Engineering** HotelTonight - San Francisco Bay... X

[Feedback | See more »](#)

Groups You May Like

 **LinkedIn Groups Product Forum** [Join - Other...](#) X

 **RDataMining** [Join - Professional Group](#) X

 **Data Mining, Statistics, and Data Visualization** [Join - Professional Group](#) X

[Feedback | See more »](#)

What is a Recommender System?

A Recommender selects a product that if acquired by the “buyer” maximizes value of both “buyer” and “seller” at a given point in time

Lesson 1 : Recommendations must make strategic sense...

- Conflicts of interest between ‘buyers’
 - E.g. job posters vs. job seekers

Director, Data Scientist
PayPal - San Jose, CA (San Francisco Bay Area)



Job Description

PayPal is the faster, safer way to pay and get paid. The service allows members to send money without sharing financial information, with the flexibility to pay using their account balances, bank accounts, credit cards or promotional financing. With nearly 100 million active accounts in 190 markets and 19 currencies around the world, PayPal enables global commerce. Putting customers' needs first, PayPal is redefining the payments category through product, marketing, and service delivery innovation. We're looking for engineers, applied scientists, and researchers with experience in implementing data mining and machine learning systems to help build innovative products all of PayPal and beyond. This is a new team under the guidance of the chief scientist at PayPal that will leverage all data throughout eBay Inc. – from creating new products, technologies and platforms, all the way to supplement existing products for various use cases. We hope to build a world-class data science team that will have direct impact on

Apply on Company Website

Save job Share job   

Stop following ▾

Posted By

Sachi Yokota 2nd
Recruiter, Marketing & Strategy at
PayPal
[Send InMail](#)

1 of your connections can refer you to Sachi:

John Reilly 1st

- What is the (long-term) value of an action?
 - How do we compare an invitation from a job application from a group join or the reading of a news?



Ingredients of a Recommender System

A Recommender processes information and transforms it into actionable knowledge

Data
(Feature
Engineering)

Algorithms

Business Logic
and Analytics

User Experience

Lesson 2: User Experience Matters Most

1. Understand **user intent**
→ Define, model, leverage
2. Be in the **user flow**
3. Optimize **location** on the page
4. Set **right** expectations (“You May...”)
5. **Explain** recommendations
6. **Interact** with the user
7. Leverage **Social Referral**

User Intent, User Flow, Location, Message

Job recommendation use cases

| User experience | Optimization | Impact on job application rate |
|-------------------------|--|--------------------------------|
| User intent / location | Homepage personalization | 2.5X |
| User flow / user intent | Before vs. after having applied to a job | 7X |
| User flow | LinkedIn homepage vs. Jobs homepage | 10X |
| Location | Center rail vs. right rail | 5X |
| Message | Followers vs. leaders | 2X |

Applied Researcher / Data Scientist: Data Analysis, Data Mining and Machine Learning
[LinkedIn](#) • Mountain View, CA (San Francisco Bay Area)



[Apply with LinkedIn](#)

[Apply with Resume](#)

Save job Share job Stop following

Posted By Shikha Mathur Technical Recruiter/Sr. Source at LinkedIn (smathur@linkedin.com) send email

59 of your connections can refer you to Shikha:

- Hillary Taubman-Dye Recruiting Manager at ConnectTV send message
- Steven Stegman Research Scientist and Sr. Product Manager Request Introduction to Shikha
- Daniel Tunkelang Principal Data Scientist at LinkedIn Request Introduction to Shikha

View all 59 people

Salary Range for Jobs Like This

| USD | Salary | Bonus | Total Cash |
|-----------------|--------|--------|------------|
| 55,700 | 7,680 | 11,170 | 74,550 |
| 25th Percentile | 40,800 | 6,600 | 47,950 |
| 50th Percentile | 60,600 | 8,800 | 69,250 |
| 75th Percentile | 75,800 | 11,200 | 87,000 |

PayScale

Job Description

The engineering culture at LinkedIn is based on building and integrating cutting-edge technologies while smoothing the user interaction and experience. Our engineers constantly raise the bar for excellence motivating each other to tackle challenges and take intelligent risks. The industry is moving fast and our engineers are right there with it.

About Data Analysis, Data Mining and Machine Learning

We're looking for an Applied Researcher / Scientist to work on innovative semi-supervised learning graph dataset. Your role is to lead a small research team that has both domain expertise in data mining, information retrieval or machine learning as a strong software orientation. It offers the unique opportunity to conduct applied research and have a huge end-to-end impact on key LinkedIn product initiatives, including search refinement, ad targeting, information extraction, and recommendations.

Desired Skills & Experience

- You are a computer science champ with strong coding ability and familiarity with Java, C++ or any other OOP language.
- You have experience with data mining toolkits like Weka, Mahout, R, NLTK, etc.
- You have experience with Hadoop or other MapReduce paradigms.
- You're familiar with information retrieval libraries like Lucene / SOLR.
- You have a solid understanding of data structures and examples for statistical modeling, data mining, recommendation or search relevance solutions.
- You have a PhD in statistics, machine learning, quantitative discipline.
- Abs. MSc. or PhD in statistics, machine learning, quantitative discipline
- You enjoy working in a fast paced, test-driven, collaborative and iterative programming environment.
- You enjoy using data to drive products.

Company Description

Founded in 2003, LinkedIn connects the world's professionals to make them more productive and successful. With more than 135 million members worldwide, including executives from every Fortune 500 company, LinkedIn is the world's largest professional network online. The company has delivered significant value to its members, helping them find better jobs, connect with colleagues, and hire sources. Headquartered in Silicon Valley, LinkedIn has offices in 22 countries across the globe.

Additional Information

Posted: November 28, 2011
 Type: Full-time
 Experience: Mid-Senior level
 Functions: Engineering, Research, Science
 Industries: Internet, Information Technology and Services, Computer Software
 Job ID: 1936094

[Apply with LinkedIn](#) [Apply with Resume](#)

People Who Viewed This Job Also Viewed

- Data Scientist at LivingSocial
- Research Scientist at Yahoo!
- Data Mining Consultant at IBM
- Software Development Engineer at LinkedIn
- Senior Algorithmic Software Engineer (C/C++) / Classification Algorithm / Image Processing / Machine Learning at KLA-Tencor
- Artificial Intelligence and Reasoning Engineer - Corporate R&D at Qualcomm
- Machine Learning Software Engineer, YouTube - Mountain View at Google
- Data Scientist (PL) at LinkedIn

Search More Jobs

- Jobs at LinkedIn
- Engineering, Research, Science Jobs
- Mid-Senior Level - Engineering, Research, Science Jobs
- Engineering, Research, Science Jobs in San Francisco Bay Area
- Internet, Information Technology and Services, Computer Software Jobs
- Mid-Senior level - Internet, Information Technology and Services, Computer Software Jobs
- Internet, Information Technology and Services, Computer Software Jobs in San Francisco Bay Area

Item based collaborative filtering:
 → Follower audience

People Who Viewed This Job Also Viewed

- Data Scientist at LivingSocial
- Research Scientist at Yahoo!
- Data Mining Consultant at IBM
- Software Development Engineer at LinkedIn
- Senior Algorithmic Software Engineer (C/C++) / Classification Algorithm / Image Processing / Machine Learning at KLA-Tencor
- Artificial Intelligence and Reasoning Engineer - Corporate R&D at Qualcomm
- Machine Learning Software Engineer, YouTube - Mountain View at Google
- Data Scientist (PL) at LinkedIn

Search More Jobs

- Jobs at LinkedIn
- Engineering, Research, Science Jobs
- Mid-Senior Level - Engineering, Research, Science Jobs
- Engineering, Research, Science Jobs in San Francisco Bay Area
- Internet, Information Technology and Services, Computer Software Jobs
- Mid-Senior level - Internet, Information Technology and Services, Computer Software Jobs
- Internet, Information Technology and Services, Computer Software Jobs in San Francisco Bay Area

Similar Jobs

- Data Scientist at LivingSocial - Washington D.C., Metro Area
- Graduate Summer Internship in Machine Learning at Intel - Bellevue, WA
- Software Engineer, Machine Learning at Olay - Mountain View, CA
- Data Scientist at Venafi - Cambridge, MA
- Principal Software Engineer at LinkedIn - Mountain View, CA

view all 9 similar jobs

Content based:
 → Leader audience

User Intent Modeling

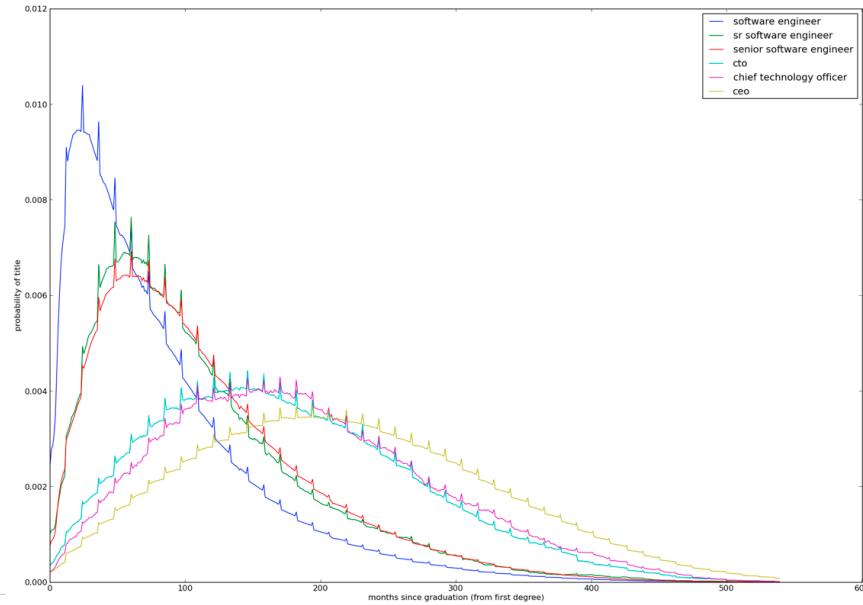
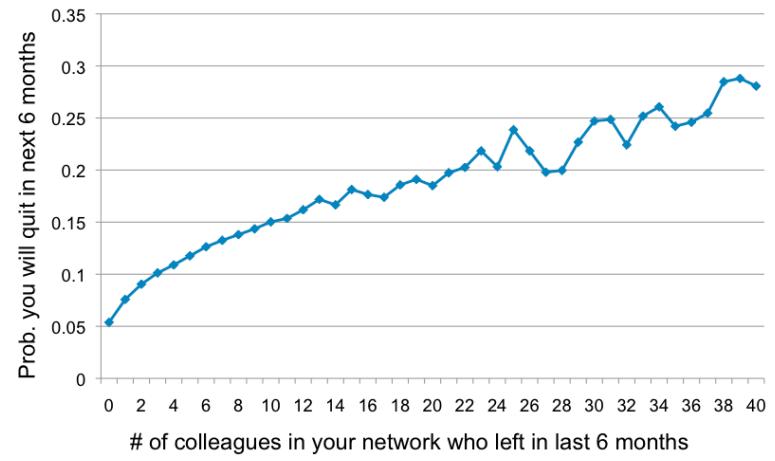
1. Intent labeling

- Job seeker, recruiter, outbound professional, content consumer, content creator, networker, profile curator

2. Build normalized propensity model for each intent

Job Seeker Intent

- Features
 - Behavior: job searches, views & applications, job related email replies, has a job seeker sub, ...
 - Social graphs: # of colleagues in network who recently left, ...
 - Content: title, industry, anniversary date, ...
- Propensity Score
 - Parametric accelerated failure time survival model
 - $\log T_i = \sum_k \beta_k x_{ik} + \sigma \varepsilon_i$
 - Score: P(switch job next month)
- Evaluation
 - Gold standard
 - Directional validation



Lesson 3: Recommendations often cater to multiple competing objectives

Groups You May Like

- Bayesian Belief Networks with BayesiaLab
Join - Professional Group
- ACM SIGKDD
Join - Professional Group
- Sentiment Analysis Symposium
Join - Conference Group

[Feedback | See more »](#)

Suggest relevant groups ...
that one is more likely to
participate in

TalentMatch
(Top 24 matches of a posted job for sale)

Home Profile Connections Groups Jobs Inbox 3 More Jobs Advanced

Your Job has been posted and is available for public viewing. Share job with your connections View job

Show: Active

Reach out to people who match your job description for: Director of Product Management

| Profile Picture | Name | Job Title | Company | Location | Score | Match | Action Buttons |
|-----------------|------------------|------------------------------|------------------------------|------------------------|-------|-------|--|
| | Susan Roberts | Product Marketing Manager | Cisco Systems | San Francisco Bay Area | 9.7 | Match | Save profile Send InMail |
| | Robert Cotton | Senior Product Manager | WebEx | San Francisco Bay Area | 9.5 | Match | Save profile Send InMail |
| | Rebecca Casewell | Director, Product Management | Hilton Worldwide | San Francisco Bay Area | 9.5 | Match | Save profile Send InMail |
| | Robert Stevens | Senior Product Manager | Dimension Data | San Francisco Bay Area | 9.4 | Match | Save profile Send InMail |
| | Jennifer Roberts | Director, Product Management | CFPM Braxis | San Francisco Bay Area | 9.2 | Match | Save profile Send InMail |
| | Steven Reynolds | Senior Manager | Presidio Networked Solutions | San Francisco Bay Area | 9.2 | Match | Save profile Send InMail |

Suggest skilled candidates ...
who will likely respond to hiring
managers inquiries

Semantic + engagement objectives

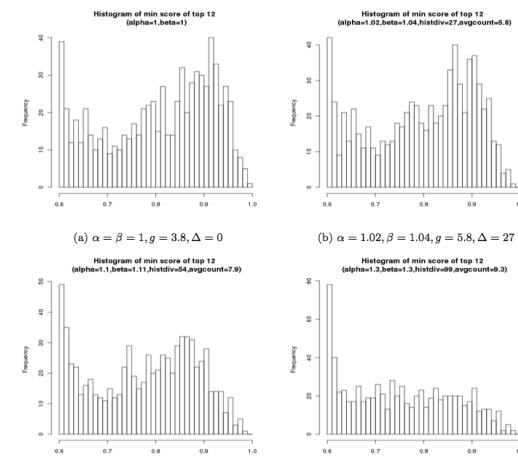
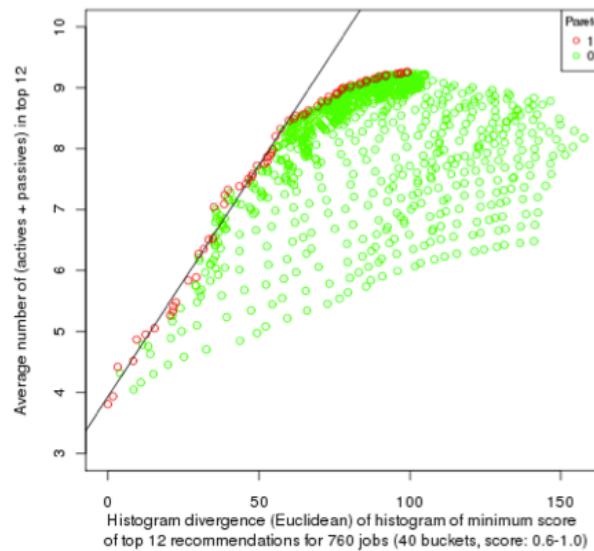
Multiple Objective Optimization

Constraint optimization problem

1. Rank top $K' > K$ results wrt to primary objective (e.g. relevance)
2. Perturb the ranking with a parametric function which leads to the inclusion of secondary objective(s) (e.g. engagement)
 - Measure the perturbation using a distance function wrt primary objective
 - Create a framework to quantify the tradeoff between the objectives

TalentMatch Use Case I

- Proxy for likelihood to answer hiring managers inquiries:
 - Job seekers are 16X more likely to answer than non-seekers
- Increase % of job seekers in TalentMatch results
- Control for TalentMatch booking rate and sent emails rate



Distribution of min TalentMatch scores over 1 month of jobs posted on LinkedIn

TalentMatch Use Case II

Table 2: InMail response rate ratio of treatment group over control group. Data collected over a period of four weeks around May/27/2012-June/17/2012.

| Treatment | Treatment/Control | 95% CI |
|-------------------------|-------------------|-------------|
| $\alpha = \beta = 1.15$ | 1.42 | 0.95 - 2.01 |
| $\alpha = \beta = 1.07$ | 1.22 | 0.94 - 1.59 |

Table 3: Booking rate ratio of treatment group over control group. Data collected over a period of four weeks around May/27/2012-June/17/2012.

| Treatment | Treatment/Control | 95% CI |
|-------------------------|-------------------|-------------|
| $\alpha = \beta = 1.15$ | 0.96 | 0.81 - 1.15 |
| $\alpha = \beta = 1.07$ | 1.00 | 0.85 - 1.20 |

Table 4: Inmails/booking ratio of treatment group over control group. For perspective, the control group has approximately 2 InMails/booking. Data collected over a period of four weeks around May/27/2012-June/17/2012.

| Treatment | Treatment/Control | 95% CI |
|-------------------------|-------------------|-------------|
| $\alpha = \beta = 1.15$ | 1.25 | 1.07 - 1.51 |
| $\alpha = \beta = 1.07$ | 1.31 | 1.11 - 1.57 |

Data Sources for Feature Engineering

Content

Applied Researcher / Data Scientist: Data Analysis, Data Mining and Machine Learning
LinkedIn - Mountain View, CA (San Francisco Bay Area)

Job Description
The engineering culture at LinkedIn is based on building and integrating cutting-edge technologies while encouraging innovation and expansion. Our engineers constantly raise the bar for themselves, collaborating with each other to tackle challenges and take intelligent risks. The industry is moving fast and our engineers are right there with it.

About Data Analysis, Data Mining and Machine Learning
We're looking for an Applied Researcher / Scientist to work on our massive semi-structured text and graph datasets. This role is ideal for a senior researcher who has both domain expertise in data mining, information retrieval, or machine learning and a strong systems orientation. It offers the unique opportunity to work on challenging problems in a fast-paced environment, and to contribute to a wide range of product initiatives, including search relevance, ad targeting, information extraction, and recommendations.

Desired Skills & Experience

- You are a computer science champ with strong coding ability and familiarity with Java, C++ or any other programming language.
- You have worked with data mining toolkits like Weka, Mahout, NLP Toolkit, etc.
- You have experience with Hadoop or other MapReduce paradigm.
- You have experience with distributed systems like Google's GFS, MapReduce, BigTable, etc.
- You're interested in how Auto-recommends billions of users for the statistical modeling, data mining, recommendation, and personalization space.
- You have put together a portfolio of work.
- You drive in.

Christian Posse
Principal Data Scientist
LinkedIn
San Francisco Bay Area

Christian Posse
Smaller Magnetic Materials
Nanotechnology
Findings from research could now power more efficient
and effective
driving solutions.
Like · Comment · Share · See activity · 4 days ago

Current Principal Data Scientist
Past Director of Technology & Director, Development - Data Corporation
Sr. Research Scientist at see all
Education Stanford University
Swiss Federal Institute of Technology Zurich
Swiss Federal Institute of Technology Lausanne
Recommendations 9 people have recommended
Connections 500+ connections
Cposse Twitter http://www.linkedin.com/in/cposse
Public Profile Share Print

Summary
Fifteen years of expertise in the research, analysis, design, development cost-effective, high-performance technology solutions to meet challenging needs in the area of predictive analysis, visual analytics, statistics, machine learning, mining, and data mining. Strong background in scientific computing, system requirements, and conceptual design in all aspects of project life cycle. Conceptual design through effective management of resources and management experience that produce as

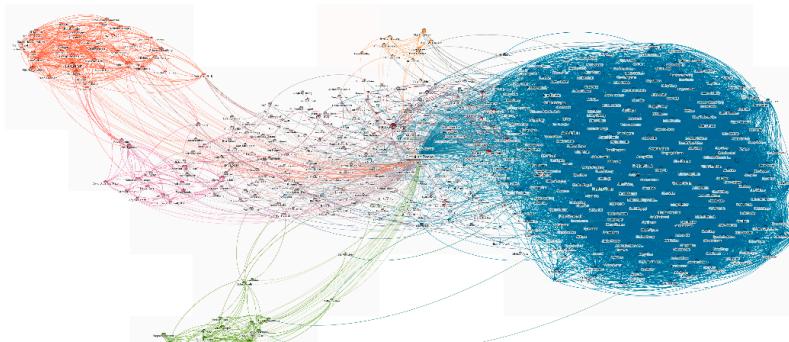
Specialties
Product Development, Product Manager, Social Computing, Software Development, Multivariate Testing, Data Mining, Model Learning, Web 2.0, Enterprise 2.0, Text & Language Processing, Information Fusion Intelligence

Experience
Principal Data Scientist and Product LinkedIn
July 2009 - Present (2 years 7 months) · M
Design and develop from ground up data infrastructure to the user experience, wor

1. Highly scale real-time content-based recommendation system for LinkedIn
2. Personalized recommendations: "People You May Want to Follow", "Jobs For Your Web Ads", "Events You May Be Interested in"
3. Similar recommendations for jobs, groups, companies, industries, profiles and events
4. Ranking and filtering products: "TalentMatch", "Referral Engine",
5. Job Agent Sourcing: Ad targeting Engine
6. Crowd-based recommendations: "Related Searches", "Viewers of This Profile Also Viewed", "People Who Viewed This Job Also Viewed", "Professionals That Recommended This Product Also Recommended People Who Viewed This Company Also Viewed", "People Also Bought These Products"

In the press: <http://slurp.com/lnmpaa>, <http://slurp.com/lnq2>, <http://slurp.com/p6oy>.

Social Graphs



Behavior

Your Groups (48) Reorder »

Visual Analytics
Those of you who present financial models - Do you use tomato charts?

The R Project for Statistical Computing
Those in Tahrir turned himself in death in December 2010 at protest in the treatment by police, pro-democracy revolution have erupted across the Middle East. Our interactive timeline tracks.

Data Mining, Statistics, and Data Visualization
George legend 1 day ago Jeffrey great presentation, we are now in 2012, is there a conclusion you can draw as to what will the Arab culture do?

Online Experimentation Professionals
TOP INDUSTRIES



Lesson 4: The Unreasonable Effectiveness of Big Data

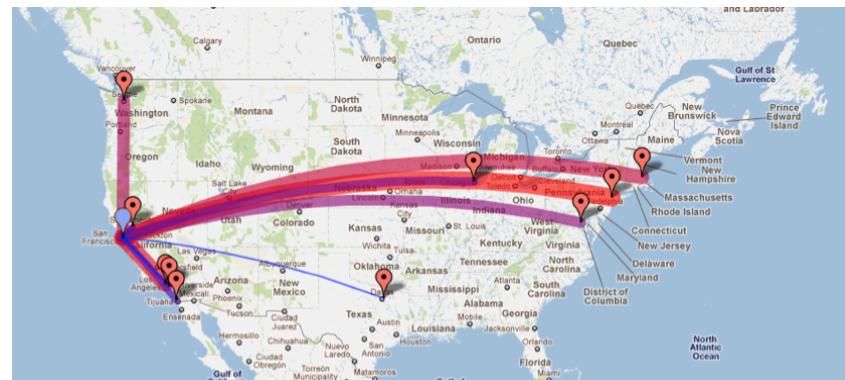
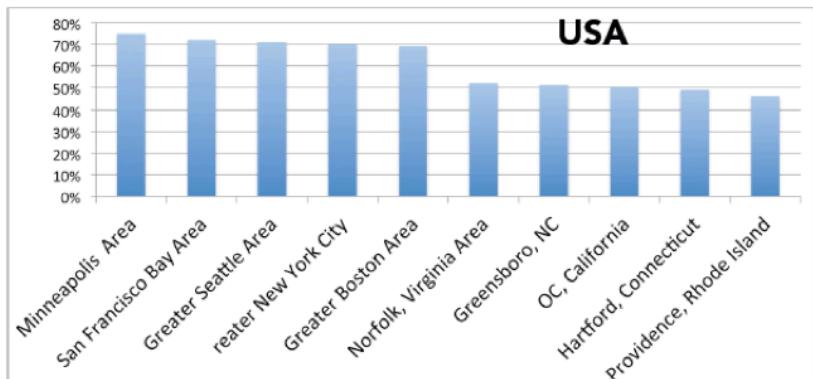
- Data jujitsu: slice and dice (smartly), then count
- Rich features engineered from profiles jujitsu
 - Job tenure distributions
 - Job transition probabilities
 - Related titles, industries, companies
 - Profile & job seniority
 - Impact on job recommendations:
 - 25% lift in views, viewers, applications, applicants
 - 90% drop in negative feedback

Lesson 4: The Unreasonable Effectiveness of Big Data (cont.)

- Region stickiness
- Related regions
 - Impact on job recommendations: 20% lift in views, viewers, applications, applicants
- Individuals propensity to migrate

Most vs Least sticky regions*

*what percentage of people stay in the same region when switching companies



Lesson 5: Data Labeling Can Be Daunting

- Historical data
 - Similar objective
 - Unrelated processes, e.g., same session search selection
 - reduce presentation bias, position bias
 - What about intent bias?
- Random suggestions
 - Great with ads, company follows
 - Not for products with high cost of bad recommendations
 - jobs, alumni groups, ...
 - Not for similar recommendations
- Crowdsourced manual labeling
 - Very challenging
 - Pairwise comparison more suited than absolute rating → higher cost
 - Expert-sourced manual labeling

Lesson 6: Measure Everything

1. Implicit user feedback

- E.g. impressions, immediate actions (clicks), secondary actions
- Understand and optimize flows, e.g.,
 - Impact of 'See more' link and landing page
 - Conversion rate of a job view into a job application from various channels
 - E.g. Homepage vs email vs mobile

JOBs YOU MAY BE INTERESTED IN

The image shows a screenshot of the LinkedIn 'Jobs You May Be Interested In' section. It displays four job recommendations with small profile icons, job titles, companies, and locations. Below the recommendations is a red rectangular box highlighting the 'See more »' button.

| Job Title | Company | Location |
|--|--------------------------|-----------------------|
| Senior Campaign Development Manager | Ads Direct Media - We... | |
| Product Manager, Big Data and... | RMS | San Francisco Bay ... |
| Senior Unstructured Data Platform... | Recommind | San Francisc... |
| Product Manager, Big Data Visualization... | Google | Mountain View, CA |

Feedback See more »

Lesson 6: Measure Everything (cont.)

2. Explicit user feedback

- Understand usefulness of the recommendation with unequal depth
 - Text based A/B test!
- Help prioritize future improvements
- Reveal unexpected complexities
 - E.g. 'Local' means different things for different members
- Prevent misinterpretation of implicit user feedback!

JOBs YOU MAY BE INTERESTED IN

 **Senior Campaign Development Manager** Sponsored
Ads Direct Media - We...


 **Product Manager, Big Data and...** ×
RMS - San Francisco Bay ...


 **Senior Unstructured Data Platform...** ×
Recommind - San Francisc...


 **Product Manager, Big Data Visualization...** ×
Google - Mountain View, CA


[Feedback](#) | [See more »](#)

From: Joseph Hicks-Malloy <jh041177@yahoo.com>
Date: August 12, 2012 5:31:32 PM PDT
Subject: [Site Feedback] Job Recommendations Feedback
To: LinkedIn Feedback <suggestions@linkedin.com>

The following comment was submitted on the LinkedIn website on August 13, 2012 12:31 AM by Joseph Hicks-Malloy:
Your job suggestions are very helpful keep up the great work

From: Steve Lane <Steve.Lane850@gmail.com>
Date: August 12, 2012 6:00:15 PM PDT
Subject: [Site Feedback] Job Recommendations Feedback
To: LinkedIn Feedback <suggestions@linkedin.com>

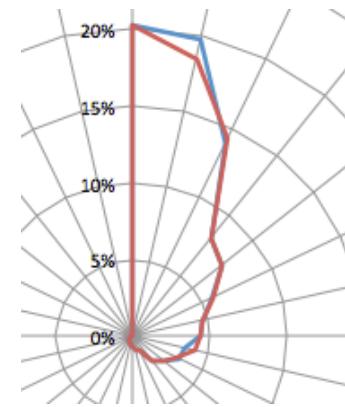
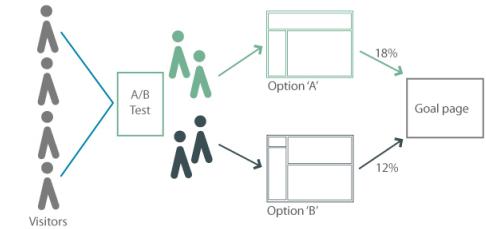
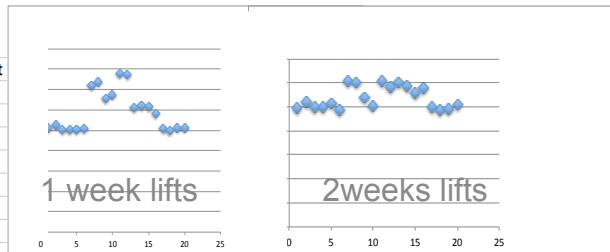
The following comment was submitted on the LinkedIn website on August 13, 2012 1:00 AM by Steve Lane:
Need jobs in memphis tn.
Steve

Lesson 7: Beware of some A/B testing pitfalls

1. Novelty effect

- E.g., new job recommendation algorithms have week-long novelty effect that shows lifts twice the stationary (real) one

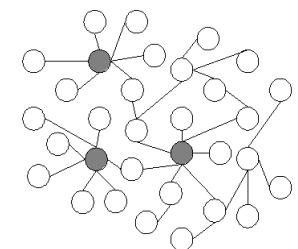
| Phase 3 | | |
|---------------------------|----------------|-------------------|
| Algorithmid in production | Explanation | PRODUCTION bucket |
| 4 | 4001 + MinMax | 0-34 |
| 5006 | 5005 + TBS-0.3 | 35-44 |
| TBD | 5005 + TBS-0.2 | 45-54 |
| TBD | 4001 + TBS-0.2 | 55-64 |
| TBD | 4001 + TBS-0.3 | 65-74 |
| TBD | 4001 + TBS-0.1 | 75-84 |
| 4 | 4001 + MinMax | 85-99 |



2. Cannibalization

- Zero-sum game or real lift?

3. Random sampling destroys network effect



Lesson 8: One Unified Platform, One API

- Scaling innovation
 - Cross-leverage improvements between products
 - Shared knowledgebase
- Maintainability
 - Production serving and tracking
 - Infrastructure for complete upgrades
- Performance
 - Billions of sub second computations

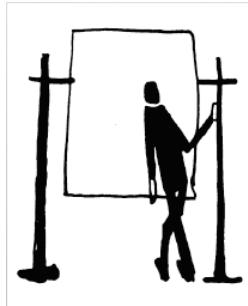
Open Source Technologies



Zoie



Bobo



Kafka



Voldemort



Conclusion

- Social/professional networks are a new frontier for recommender systems
- Still many open questions:
 - How do we define and measure engagement?
 - What is the utility of a recommendation for the member?
 - What is the value of a recommendation for the network?
 - How do we reconcile utility and value when they conflict?
 - How do we network A/B test without tears?
 - ...
- Learn as you go
 - Track everything and invest in forensics analytics
 - Breadth – understand holistic impact
 - Depth – understand flows

References

- M. Rodriguez, C. Posse and E. Zhang. 2012. Multiple Objective Optimization in Recommendation Systems. To appear in *Proceedings of the Sixth ACM conference on Recommender systems* (RecSys '12)
- M. Amin, B. Yan, S. Sriram, A. Bhasin and C. Posse. 2012. Social Referral : Using Network Connections to Deliver Recommendations. To appear in *Proceedings of the Sixth ACM conference on Recommender systems* (RecSys '12)
- A. Reda, Y. Park, M. Tiwari, C. Posse and S. Shah. 2012. Metaphor: Related Search Recommendations on a Social Network. 2012. To appear in *Proceedings of the 21st International Conference on Information and Knowledge Management* (CIKM '12)
- B. Yan, I. Bajaj and A. Bhasin. 2011. Entity Resolution Using Social Graphs for Business Applications. In *Proceedings of the International Conference on Advances in Social Network Analysis and Mining* (ASONAM '11)