Overview

Use this section to provide a brief overview of the reasons for placement at the proposed job level. Introduce the candidate (i.e. role, organization, and specialty) and highlight the primary reasons for promotion, including a summary of the candidate’s key technical contributions and their impact on the business/industry/customers.

Lakshminarayan is an Engineering Fellow at Teradata Labs. He is a researcher, technologist, developer, inventor, consultant, educator, evangelist, an author, and dedicated to the advancement, application, and dissemination of statistics. His core expertise is mathematical statistics strongly supported by Applied Mathematics, Computational Statistics, Statistical Signal Processing, Data Mining, Machine Learning, Scientific Computing, Industrial Mathematics and Statistical Process control and Design of Experiments. The expertise is applied to many application areas ranging digital marketing, marketing optimization, e-commerce, energy, finance, query optimization, statistical and mathematical algorithms, cyber-security, healthcare, real estate, and semiconductor manufacturing. In each of the areas, he made impactful contributions. As a technologist and developer, he wrote algorithms that power marketing optimization applications, query optimizers in databases, customer segmentation in digital marketing, and sales pipeline analytics to forecast revenues. As a researcher he developed algorithms for anomaly detection in data centers, detecting turbulence in oil flows (energy), resource utilization in intensive care units (ICU), algorithms for anomalies in bio-signals (wearable computing). These research efforts were developed as part of proof of concepts (PoC), and demonstrators. He holds 24 granted patents, 3 allowed patents, a defensive publication, and 20 more pending. He has published many peer reviewed papers in journals and international conferences. As a consultant he provided consulting expertise across businesses in banking, real estate, medical claims, legal, industrial design, and auto financing. As an educator and evangelist, he speaks at symposia, colloquia, conferences, local chapters of IEEE, the American statistical Association, and teaches courses on “Statistics and Big Data” and Linear model theory as adjunct assistant professor at the University of Texas at Austin. He also is writing a book (the CRC Press-Chapman-Hall) on predictive analytics for graduate students and industrial practitioners and serves as a reviewer and conference program committee member. He mentors and mentors many students and colleagues, many of whom are well placed in APPLE, HP, Facebook, SAS Institute, Texas Instruments, and AT &T to name a few.

Significant technical achievements

**Parallel Computing**

Teradata Labs is focused on large-scale machine learning. As data sets are growing in size, exemplified by analyses of text data, genomic data, customer transactions data, speech, and video, ML algorithms have to be redesigned and rewritten for parallel computation.

**My Contribution and Results**

Lakshminarayan as the chief data scientist has been crucial to rewrite the generalized linear model (GLM) package for linear regression with Gaussian errors and logistic regression with and regularization. He also contributed to the XGBoost algorithm, as well as Latent Dirichlet Allocation (LDA), and Support vector machine (SVM), Data preparation functions to name a few. He collaborates with customers to develop new techniques for their unique problems and the new methods are incorporated into the TERADATA machine learning engine. He contributes to the Teradata intellectual property portfolio through patent applications. He serves as a trainer conducting workshops on Machine learning and Statistics. He serves as an examiner for patent application approvals.

**Results**

1. Machine learning and statistical algorithms listed above implemented in the commercially available Teradata Machine Learning Engine.
2. *System and methods for acceleration of machine learning functions,* Patent pending
3. *Information Theoretic Models for Statistical Classification, Patent pending*

**Sales Pipeline Analytics**

**Problem Statement**: HP software uses salesforce.com data to track deals in the sales pipeline. Accurate prediction of “open” deals in the pipeline is desirable so resources (sales representative’s workload) allocation by deal can be streamlined to meet sales targets. It is therefore desirable to build a program to predict the likelihood of “deals closing” accurately. So revenue can be forecasted accurately and resources can be adjusted to help deals to close successfully. Current methodology for pipeline forecasting is merely to use intuition and subjective judgment.

**My contribution**: Worked with the global pre-sales team to scope the problem, identify relevant data for modeling and build a variety of statistical/machine learning tools. The output of the models is converted into a report to provide a snapshot of deals’ status for planning and closing to the managers and sales representatives.

1. **Results of Modeling**: Our statistical (machine learning) model predicts the likelihood of deals closing with ~90% accuracy with a 10% false positive rate.
2. The output is a simple spreadsheet that ranks all open deals by likelihood of closing together with summary statistics about statistically significant attributes that influenced the deals’ outcomes.
3. Models were built with actual (raw) sales pipeline data from Sales Force
4. The model output-spreadsheet, is now used to track open deals in Q4-2015, Q1-2016.

*Commenting on the collaboration, John Margaglione, Vice President of sales at HP software, wrote, “I engaged Lakshminarayan based on my knowledge of his past work with HP Labs, and based on the regular conversations we have been having around the subject of pipeline management over the past year. Despite his commitments to work and his publisher, he agreed to assist me with the project. The results so far have been spectacular.*

*Lakshminarayan has a consultative approach that quickly led to capturing the most relevant aspects of the project, and uncovered several missed opportunities in previous work. He also imparts knowledge at a rapid pace: I always feel a little smarter after every encounter. He is very easy to work with, and does not waste time. His grasp of the available mathematical models, and his impartiality in selecting the right models, resulted in breakthroughs that the BU has not been able to achieve, even with a dedicated statistician”*

**Digital Marketing**

While segmentation technologies abound, marketers overwhelmingly complain that those technologies do not produce segments aligned with key performance indicators (KPI) while being actionable in real-time. Generally, the solution is successful if a lift in a key performance indicator (KPI) is obtained when campaign resources are diverted from the general population to that segment and the results are “actionable” in that explicit characteristics found can be used directly, to filter and promote to consumers, in their digital channel’s real-time infrastructure.

My contribution:

The cornerstone of our solution consists of two parts: First is the construction of a “contextual space” for each attribute. This space inherits a topology from some quantifier of a context, in this case the response variable (KPI) assigning a point in the space for each value of each attribute such that distances are measured relative to the context. This helps to link the attribute to the KPI introducing the notion of relevance. In the next step, the contextualized data is the input to a statistical clustering algorithm to generate homogeneous groups of entities (customers). The quality of the clustering in typical segmentation systems rely on model identification statistics such as Akaike information criterion (AIC), and Bayesian Information Criterion (BIC). Sadly these statistics produce too many clusters to be actionable to a marketer. My specific contributions are:

1. Contextualizing the data by a non-linear transformation
2. Developed an algorithm that automates the generation of optimal number of clusters based on signal to noise ratios (SNR). The clusters produced are stable, homogeneous, and actionable for digital marketing.
3. Developed heuristics for stopping rules. It is an algorithm that iterates through various values of K (K greater than equal to 2), and measures the change on SNR form iteration to iteration. Based on a properly chosen threshold, it informs the selection of optimal number of clusters.

**Results**

1. Our behavioral analysis solution is a key component of the new multi-channel analytics product from Marketing Optimization, the HP Digital Marketing Hub (DMH), which was launched in October 2013 for general availability. This is an industry leading solution with a wide adoption.
2. “Segments of Contacts,” United States Patent, *10,565,603*

*Commenting on the solution Lakshminarayan developed, Roger Brooks, Chief Technologist and Distinguished Technologist at HP said, “The Digital Marketing Hub” product for behavioral analytics could not be released for general availability without Lakshminarayan’s key contribution of stopping rules to determine optimal number of clusters and testing on customer datasets…”*

**Marketing Optimization**

HP Corporate Marketing needed a systematic and reliable way of assessing marketing effectiveness and optimally allocating investments among marketing vehicles given a fixed budget. As marketing budgets are planned, the methods and tools we developed assist managers to assess allocation scenarios and determine investment strategies to maximize revenue. The novelty of our approach is the innovative integration of multiple technical disciplines to produce a unified framework and create an easy to implement end-user solution for revenue/profit maximization. The methodology is flexible to be adopted in other areas where optimization is necessary.

**Our Solution**

The goal of Marketing Optimization (Return on Marketing Investment-RoMI) is to determine investments in marketing vehicles to maximize revenue or margin. The solution utilizes principles of *market mix modeling* (MMM). Market mix modeling consists of econometrics, Delphi method for incorporating expert knowledge and mathematical optimization. Econometrics is used to determine the effect of marketing vehicles on revenue. The Delphi method is applied to obtain expert opinion regarding additional factors that may affect return on investment. The expert analysis is blended into the econometric equation to improve the predictive capacity of the model. Finally optimization is invoked to find a mix of marketing investments that yield highest revenues given a budget.

*Commenting on the program, The Vice president of HP Marketing, Nigel Ball said, “The ROMI tool will give us common metrics to take the guess-work out of planning, make data-driven decisions, and clearly link marketing strategies to business results…..The models give precise data on what mix of marketing drivers-advertising, events, direct marketing, etc.--delivers the best results for each TSG product line…… ”*

**My Contribution**

1. Lakshminarayan lead the effort to first interview outside vendors (Accenture, Marketics, other vendors) to choose a third party to pilot initial Marketing Optimization.
2. Lakshminarayan lead the HP side to engage with Accenture on the Technical aspects.
3. When HP owned the responsibility to productize RoMI, Lakshminarayan developed the basic econometric, quantification of qualitative data, and optimization tools.
4. It is noteworthy to mention that Lakshminarayan not only created the underlying methodology, but also wrote the preliminary prototype of the RoMI system. The systems and tools developed is key intellectual property of HP.
5. Lakshminarayan guided a team of 5 engineers in Bengaluru, India to productize and duplicate the models in Germany, France, Netherlands, Canada, Spain, France, and Mexico.

**Results**

|  |
| --- |
| 1. The program is implemented in Australia, Canada, Netherlands, and Korea and near-term implementation schedule includes Spain, Germany Mexico, and France. RoMI recommendations in Canada in 2006 yielded positive results. The average ROI increased from 7.5x to 10.5x. By implementing according to recommended optimal marketing investments, a 36% increase totaling **15 million USD** in incremental revenue was achieved. Including Australia and Korea, the incremental revenue is approximately **30 Million USD**. 2. *A system for accurately predicting revenues given a portfolio of investments, 8,180,693* 3. *System and method for optimizing financial performance generated by marketing investments under budget constraints, 8,027,897* 4. *A system and Method for incorporating expert opinion in statistical models, 8,180,694* |

**Update Statistics for Query Optimization**

Query optimizers in Database management system (DBMS) have to produce efficient plans for executing queries. The optimizer relies heavily on statistics pertaining to the data in the tables known as *update statistics*. A key problem is to accurately estimate the number of unique values in database columns (because this affects the number of rows accessed or processed). The original estimator was based on simplistic assumptions and gross approximations and believed to affect query planning and execution at customer sites and numerous dissatisfied customers. Lakshminarayan was approached by NonStop engineering to research the problem in 2007. His investigations led to designing an estimator that was much more accurate. Unique values estimation in a dataset is sensitive to degree of asymmetry in the distribution of the original observations. It is known as skewness.

The Update statistics work prolonged and continued over the years due to changing business landscape and so was revisited from time to time. The Seaquest database group kept seeking my involvement over the years until recently (2014). As the workloads kept increasing leading to Big Data regimes, HPIT was dissatisfied with the optimizer performance. So rather than refreshing the Update statistics on a regularly scheduled basis, the algorithms needed to be modified for “incremental updating” as the tables were updated with new tuples. Lakshminarayan worked with HPIT to deliver a set of algorithms for incremental update statistics (IUS). The IUS algorithms are currently used by HPIT in Query Optimization.

**My Contribution**

1. Algorithm research, development, and prototyping in SAS/MATLAB was entirely Lakshminarayan’s responsibility. He extensively studied the Statistics and Database literature to find the “Best in Class” algorithms and modified them for Query Optimization purposes.
2. Lakshminarayan tested the algorithm performance on simulated datasets as well as customer data to evaluate performance.
3. Collaborated closely with product engineers for implementation in production systems for general availability to customers

**Results**

1. The adaptive estimator delivers far more accurate performance compared to the existing estimator. At 1% sampling rate of large customer tables, the *average* improvement is 48%.
2. In production systems in NASDAQ, Wal-Mart, and other retailers
3. *System and Method for estimating unique attributes in a database, US patent, 7,428,550*
4. *Estimation of Unique Database Values, US patent, 8,549,004*
5. *Estimate the number of unique values in a list, US patent, 9,158, 815*
6. *Estimating Unique Entry Counts Using a Counting Bloom Filter, US patent, 9, 465,826.*
7. *Incrementally Updating Statistics, US Patent, 10,430, 411*

*Referring to his contributions, the Sr. Director of HP Software, Lily Jow, wrote, “I want to congratulate you for your innovation and thank you for your contribution to the Neoview product. I also appreciate the support from your management team.”*

*The then Chief Technologist of Neoview/Seaquest Rohit Jain hired Lakshminarayan by cross-charging his services to his parent cost center and said, “HPIT is very indebted to your contributions to the product that made significant business impact….”*

**Business Intelligence and Live Analytics (2010-2013)**

**Problem Statement**

At HP Labs, in the Intelligent Analytics Laboratory (IAL), there was an effort to support the Oil and Gas industry to do Live Analytics. Working with Enterprise services, we identified the need for fast, on-line algorithms to analyze of high data-rate measurements obtained from sensors during drilling. Given the ever-increasing number of sensors coupled with the increasing complexity of applications, and workload patterns, detection methods must operate in near real-time. The flow-rate of oil is not static but arrives in a streaming fashion. Temporal data varies over time. The stream is an ever-changing superposition of distributions with properties such as varying periodicities and -amplitudes.  The detection and modeling of such changes in real-time is one of the fundamental challenges in data stream processing.

**My Contribution**

1. Mine the oil flow-rates data to identify high oscillation-high amplitude, low oscillation-high-amplitude, high-oscillation, low-amplitude, and normal regimes.
2. Model turbulent (churning and slugging) phenomena in oil flows by linear and non-linear time series methods.
3. Results showed that the average false positive rate and true positive rates of ~2% and 96% respectively.
4. Further advanced the analytics operators in 2011 to detect patterns in oil and gas flow rates by the *mixture of experts*. It encompasses such diverse operators as: Short-time Fourier Transform, Dynamical Systems, Auto Regressive processes and the Integral and Fire methods.

**Results**

1. A demonstrator/proof of concept built for British Petroleum
2. *Multi Regime Detection in Streaming Data, US patent, 8,620,987*
3. *Anomaly Detection in Streaming Data, US patent, 9,218,527*
4. *Detecting Regime Change in Streaming Data, US Patent, 9,078,629*

**Data Center Operations**

**Problem Statement**

Anomalies in Data Center due to hot-spots (high memory and High CPU, other metrics) are common. The problem was from American Airlines Data Centers.

**My Contribution**

1. Developed algorithms for performing *anomaly detection and ranking* in data centers by performing statistical analysis of the various server level metrics.
2. These algorithms are meant to be applied in scenarios where a large number of metrics have to be monitored.
3. The algorithms were tested on production quality data through a pilot system that operated on the data from a Data Center belonging to an Enterprise Services customer.

**Results**

1. A prototype was built for American Airlines to monitor thousands of servers and for HP Information Technology services
2. *Anomaly Detection in Data Centers, US Patent, 8,668,620*
3. *System and Method for Ranking Anomalies, 9,141 914*

Demonstrated technical leadership

**Technology/Business Innovation:**

1. Uses innovation to increase company revenue/productivity. Examples include Sales Pipeline Analytics, Analytics for Digital Marketing, Marketing Optimization, Query Optimization in Databases, Healthcare analytics, and Internet of Things.
2. Demonstrates sound technical and business judgment: Outcomes of Lakshminarayan’s work resulted in programs for revenue optimization implemented in Germany, Canada, Korea, Spain, The Netherlands, and Mexico. Novel product features in query optimizers for efficient sampling and computing statistics, Actionable algorithms for customer segmentation. Resulting multiple (granted, pending). A wide range of Industries consult Lakshminarayan to help solve complex analytics problems.
3. Demonstrates excellence in applying new technologies to business processes or services (see C.V. in the section of Papers and Patents)
4. Possesses broad knowledge of technologies and applications: Expertise in Mathematical Statistics, Computational Statistics, Machine Learning, Signal Processing, Applied Mathematics applied to Digital Marketing, Parallel algorithms, Healthcare, Oil and Gas, Web Analytics, Sales Pipeline Analytics, Recommendation Systems, Semiconductor Manufacturing, Semiconductor yield enhancement, Banking, Database Marketing, and Customer reviews and Text Mining (See C.V. on consulting). He is sought by many organizations Institutions (IEEE, American Statistical Association, The University of Texas at Austin, Arlington to conduct seminars, workshops. Regular speaker at universities abroad (INRIA, France, University of Vienna, Imperial College, London, American University, Washington, D.C., University of North texas, The Indian Institute of Science, The Indian Institute of Technology) to name a few.
5. Owns 22 granted patents, 3 allowed patents , one defensive publication, with over 30 pending
6. Over 60 papers, Technical reports in peer-reviewed conferences, journals (See C.V)

|  |
| --- |
| **Action/Urgency Oriented:**   1. Accepts challenges 2. Uses and updates skills, technologies and new processes (See C.V. for Technical skills and breadth of areas of application)( 3. Builds, improves and advances the business 4. Made a specific, vital contribution to the business |
| Lakshminarayan was approached by NonStop engineering when he was a member of HP Labs to research the problem of computing statistics for Query Optimization in databases. His compensation was cross-charged to the database group. His investigations led to designing an estimator that was much more accurate in estimating the number of distinct values in the columns of a database table. The solution was adopted in the product and is still used by HPIT. He was granted 3 U.S. patents for his work and has several more pending. His follow-up work on “Incremental Update Statistics,” is now adopted by HPIT and he has 4 more patents pending with the United States Patent Office (See C.V. section on Patents Pending).   1. When Lakshminarayan moved from the Marketing Optimization group in Autonomy, they negotiated with Vertica Management to continue to support and contribute to the Digital Marketing Hub product over a period of 6 months by borrowing significant fraction of his time until separation to HP Inc in September. 2. The Vice President of worldwide pre-sales in Vertica approached Lakshminarayan to help build a program around Sales Pipeline Analytics. The program is underway, and we have a production-caliber algorithms to predict open deals in the pipeline. |

**People Development:**

1. Mentorship: Directed doctoral theses of PhD and Masters students, Conducts workshops and short courses around the world
2. His mentees of gone on to obtain permanent positions in HP, Apple, Facebook, SAS, Amazon, Texas Instruments, Teradata, Paypal among others
3. His students have been placed at The University of Texas at Austin, The University of Texas at Dallas, Southern Methodist University, Rice University, The University of Southern California among others.
4. Fosters environment of creativity and innovation

Demonstrated breadth and depth of knowledge

Identify the candidate’s primary technical competencies and explain how they have been demonstrated. Emphasize concrete examples, results and sustained abilities in the following areas: Breadth of knowledge, Depth of knowledge and Emerging Technologies.

|  |
| --- |
| **Personal Leadership:**   1. Lakshminarayan has strong thought leadership skills by engaging with business groups to build innovative solutions and programs with significant business impact(See section on Significant Technical achievements) 2. He initiates and influences decisions and obtains needed resources. More recently he influenced the Vertica worldwide pre-sales management to adopt advanced analytics for sales pipeline prediction. At Autonomy, Marketing Optimization his influence resulted in a comprehensive study of clickstream analytics (web path optimization) and developed a solution that is actionable in customer segmentation. As well as proposed solution for optimal sampling for A/B Testing and customer analytics. 3. Lakshminarayan promotes continuous learning via conducting seminars, teaching graduate-level courses, lectures at conferences, universities, and professional forums (IEEE, American Statistical Association). 4. Achieves national/international prominence or recognition by external professional involvement: Invited to write by a book on Big Data by the CRC Press (Chapman, and Hall), See exhibit B, Recognized as a leader by peers, Visiting professor IIT-Hyderabad, Visiting professor Indian Institute of Information Technology-Bengaluru, Adjunct Assistant Professor, UT-Austin) 5. Develops talent outside of the company via universities, public education, professional societies, and communities of interest. Joint research with University of Florida at Gainsville, University of Texas at Arlington, The University of Texas at Dallas, The University of Texas at Austin, The Indian Institute of Science, Bengaluru, Georgia Institute of Technology, Atlanta. Mentors students that led to advanced degrees (See co-advising section in C.V.). 6. Leads by example: When a complex problem is presented, he learns the existing literature, prototypes a solution (coding), and makes a passionate marketing pitch to influence the direction of research and product development. |
| * Lakshminarayan is a recognized expert in statistical theory, statistical methods, Machine Learning, Applied Mathematics, and query optimizers in databases. He has developed statistical applications in the areas of clustering and classification of structured, semi-structured, and unstructured data, methods for analyzing clickstream data, algorithms for data compression (dimensionality reduction), Signal deconvolution of Bio-Signals.  1. Lakshminarayan earned his Doctorate, Master of Science (2), and Bachelor of Science degrees, all in all in Mathematical Sciences with an emphasis in Mathematical Statistics, and Mathematical Optimization the University of Texas and Osmania University in India. 2. Lakshminarayan had a productive career in Texas Instruments as process control engineer (member technical staff), Motorola in Advanced Products Research and Development Labs (APDRL) in Motorola as a research member in the yield enhancement group, Compaq Computer Corporation, HP Labs, and Teradat Labs, and taught part-time at the University of Texas at Dallas, Department of Mathematical sciences, and at the Department of Management Sciences at the University of Texas at Arlington. And conducted academic research specializing in Statistical pattern Recognition and Generalized Linear Models.. 3. Lakshminarayan has researched complex technical problems under research grants from the Texas Instruments Semiconductor technical council, the Open Innovation Office from HP Labs, Texas Advanced Technology Research, and pending proposals at NSF. 4. He is a co-founder of Pattern Sciences Inc., an Austin based Machine learning/Artificial Intelligence start up |

Education summary

List the college degrees and other key professional/academic qualifications, including dates.

Recommendation letters

Attach recommendation letters, and include a list of the letter author(s) with their titles, organization, and company name.