DATTARAM B A

CONTACT DETAILS

E-Mail: badattaram2@gmail.com **Mob:** +91-9591982000

- I am looking for Opportunities to work on Machine Learning/Statistical Algorithms based Projects.
- Extensive experience of working with Research groups on various Data Mining/ Machine Learning Projects.
- Contributions to the domain through Technical articles, research oriented experiments to evaluate new techniques/algorithms, Patents, Co-authoring chapters in books etc.
- One year Executive Education from Indian Institute of Management, Bangalore (IIMB) on "Business Analytics and Intelligence."
- Extensive experience of working with Teams which are spread across Globally.
- Experience in customer requirement analysis, design and planning of deliverables for Product Release.
- A total of 10+ years of industrial experience.

WORK EXPERIENCE SUMMARY			
Company	**Designation**	**To – From**	**Duration**
IBM India Pvt. Ltd.	Staff software engineer	December 2005 – Till date	8 Years 4 months.
Sanyo India Pvt. Ltd	Embedded S/w engineer	January 2004 – December 2005	2 years
Yokogowa India Pvt Ltd	Trainee Engineer	September 2003 – December 2003	3months

EDUCATION			
Year	**Course**	**Specialization**	**Institute/ Board**
2012	M.Tech (4th sem)	IT	Karnataka State Open University
2012	Executive Education	Business Analytics and Intelligence	**Indian Institute of Management, Bangalore (IIMB)**
2003	B.E.	Instrumentation and Electronics	SIT,Tumkur,Karnataka
1999	XII	-	Sarvodaya PU College, Tumkur
1997	X	-	VVPHS,Tumkur

Technical Contributions

a) Patents

Generation of location based Instant ads for Groups (http://ip.com/IPCOM/000227451)	Details: The Patent proposes a new idea of delivery of Instant ads for groups of people like friends using Computer Vision.

b) Developer Articles on IBM Developer Works

1) Synchronize data with control signals in the InfoSphere Streams Time Series Toolkit.	The article was published to help the customers in understanding the synchronization mechanism for data with control signal. This was a request from the customers.
2) Real-time anomaly detection using InfoSphere Streams Time Series toolkit.	The article was published to help customers in implement the widely used Anomaly Detection application in real time. Customers from Cyber security, Healthcare use this widely.

c) DeveloperWorks This Week on YouTube

Real-time anomaly detection using InfoSphere Streams Time Series toolkit. (www.youtube.com/watch?v=dZj0pvyw-U4)	This is a video shot by devWorks This Week team. In this video, we have explained how to develop an Anomaly detection system using TimeSeries toolkits

d) Redbook

InfoSphere Streams: Accelerating Deployments Using Analytic Accelerators.	I have co-authored a chapter on TimeSeries toolkits in this redbook. The chapter 11 explains the algorithms and timeSeries models like ARIMA, Holtwinter, Kalman,Gaussian Additive Models and many Signal processing algorithms like CrossCorrelate. The chapter also dwells in to the possible applications and also provides a basic mathematical details of each algorithm.

PROFESSIONAL SYNOPSIS

Position and Responsibilities:

IBM India Pvt. Ltd. (December 2005 – Till date)

Current Position: Real Time Analytics Product Developer in Infosphere Streams product. Infosphere Streams is a Real Time Analytics Product (IBM's Real Time Product for BigData) with many built-in accelerators. I am working with IBM's Research Labs on addition and enhancement of mathematical models for the TimeSeries Toolkit Accelerator. TimeSeries Toolkit accelerator is a collection of machine learning and timeSeries algorithms.

- I'm working for TimeSeries Toolkit development in Infosphere Streams
- My role involves interacting with IBM Watson Research group for design and development of TimeSeries operators and algorithms for models like ARIMA, Holtwinters, PCA, VAR, Gaussian Additive Models (GAM) and PSAX based on Customer requirements.

Position before joining Machine Learning group: Component Lead for CPU Team.

- I was part of the Microprocessor development in Systems and Technology group.
- Extensively involved with development of System software for IBM's System Z (Mainframes) and System P (POWER series.)

Sanyo India Pvt Ltd. (January 2004 – December 2005)

- Worked as Embedded Software development Engineer. I was involved with the development of "Digital Energy Meter"
- Design and Development of TV games for Analog Televisions.

SKILL SET

Analytics	
Statistical Methods	Simple Linear Regression, Logistic Regression, Stochastic Models – Markov Chains, TimeSeries models.
Machine Learning	Hidden Markov Models (HMM's), Topic Modelling,Knowledge of SVM's, Clustering algorithms and PCA.
Statistical Tools	R, Open Bugs (for Bayes), SPSS,BigInsights, Mahout (limited experience)
Statistical Programming	R
Hardware and Software:	
Language	Assembly, C,C++, Python (Basics)
Processors/ Systems	System P and System Z Knowledge. Extensively worked on Power and Z/Arch Processor.
Programming	Knowledge of Parallel Programming, Map Reduce and Hadoop
Operating Systems	Linux
Other Interests	Artificial Intelligence, Statistical Natural Language Processing and Human Psychology.

Analytics Product Development Projects:

Project	InfoSphere Streams (Real Time Analytics)
Model	Design and Development of Principal Component Analysis(PCA) Algorithm
Details	PCA is a very widely used algorithm for dimensionality reduction. The work item involved design of PCA frontend (operator) and the algorithm. Eigen Vector and Eigen Value based PCA was developed using LAPACK library.

Project	InfoSphere Streams (Real Time Analytics)
Model	Design and Development of Control port feature for TimeSeries models like ARIMA, Holtwinter, VAR and Gaussian Additive Models (GAM)
Details	TimeSeries models is very widely used for forecasting operations. Inorder to allow the users to Monitor the models (i.e. coefficients of models like alpha, beta and Gamma in Holtwinter's) and also allow the users to load their own models, a new feature called Control Port feature was added to many of the existing operators.

Project	InfoSphere Streams (Real Time Analytics)
Model	Design and Development of Interpolate and PSAX (Piecewise Symbolic Approximation algorithms)models in TimeSeries
Details	TimeSeries data suffer from lot of noise and missing data. Inorder to allow the users to fill the missing data with predictive methods, IncrementalInterpolate operator supporting different methods was developed. This would allow customers specifically from healthcare domain to use this operator to fill the missing values. PSAX (Piecewise Symbolic Approximation algorithms) is widely used for converting timeSeries data to symbolic values. This would allow users to convert numerical data to symbols and then feed it to different algorithms which act upon symbols.

Project	InfoSphere Streams (Real Time Analytics)
Model	Design and Development of IncrementalInterpolate in TimeSeries Toolkit.
Details	TimeSeries data usually suffers from missing values especially in domains like healthcare. Before modeling the data, it is necessary to fill the missing values. IncrementalInterpolate was developed for filling the missing values. It provides many different methods like Average, last or predictive methods. Users can chose any of these methods for interpolation.

Analytics Modeling Projects:

Project	Detection of Cyber-Bullying Comments using Machine Learning Methods.
Details	Social media provides a platform for its users to express themselves. This platform is misused by many users to hurt and post unethical comments. These comments are known as Bullying comments. These Bullying comments should be detected and appropriate action needs to be taken. In the project, we used variety of machine learning techniques like K-NN, Maximum Entropy for detection and classification of cyber-bullying comments.
Tool	R
Location	In Collabration with IBM India Research Labs (IRL)

Project	Philliphines Elections 2013: Extraction of patterns and topics being discussed on Twitter
Details	Philliphines Elections was held in 2013. Social media played a key role during this event. Citizens used twitter social media to express their opinions in various forms. It was also used as a platform to discuss various national issues
Solution	Topic Modeling was used to extract top 5 topics being discussed during various stages of pre-elections. Topic Modeling provided an insight in to what the voters are expecting from the new political party.
Tool	R
Location	In Collabration with IBM India Research Labs (IRL)

Project: Protein Classification Using Kernel Methods Carried Out at IIMB

Project	Protein Classification Using Fisher Score vector (G-Protein Coupled Receptor class of Proteins.)
Source	The source of the project is the Research Paper " Using the Fisher Kernel Method to detect remote Protein Homologies" by Tommi, Mark and David.
Project Details	Proteins are represented as a series of amino acids with each amino acid represented by a alphabet. The Proteins belonging to the same family share 25% to 30% of homologous proteins. Task in this project is to detect G-Protein Coupled Receptors Proteins.

Model	1) Hidden Markov Models is used to represent the Protein Sequence Generation Process. 2) HMM's are trained using Baum Welch Algorithm. 3) Fisher score is obtained as described in the research paper. 4) SVM is trained for positive and negative examples using Fisher score vector. 5) SVM classifies the given Protein Sequence as a sub-class of GPCR or not.
Guide	Pulak Ghosh, Professor at IIMB in the Dept. of Statistics.
Tool	R

Second Project: Maximizing Revenue for the a Restaurant Chain (name of the Restaurant is with held) Carried Out at IIMB

Project	A Restaurant chain based in Bangalore, Chennai, Hyderabad and Pune.
Project Details	The Restaurant Chain has 23 branches in Bangalore with a Centralized Kitchen. The Policy of the Restaurant is to serve Fresh Food and hence the Wastage must be minimal.
Model	Predictive Analytics employed to predict each day's requirement of sub-products (like number of rotis or biryani paste) needed for each branch. Different models Linear Regression and timeSeries models was developed to predict the requirements for each day.
Tool	R
Guide	Dinesh Kumar, Professor at IIMB in the Dept. of Statistics.

M.Tech 4th Sem Project

Project	Deep Learning based NLP (ongoing work)
Details	Deep Learning is a new Machine Learning based algorithms. Deep Learning is the first step towards Noetic NLP. The projext is aimed at application of deep learning techniques for text classification or Dimensionality reduction.
Tool	R, Python
Guide	Senior Technical Person from IBM India Research Labs.

University Relations Activities

- Mentored a project on Microcontroller based "Digital Attendance" in 2006.
- Delivered a lecture on "Advanced MicroProcessor and Microarchitecture" at SIT in 2008.
- Member of the "Study Group" representing Industries at SIT, I & E dept. Role involves project guiding, Syllabus formation for Deemed University.
- Involved in BigData material preparation as part of IBM's effort towards BigData University.

Personal Details

Name	B.A.DATTARAM
Date Of Birth	13th Dec 1981
Marital Status	Married
Wife's Name	B.K.BHAGYSHREE
Address	#70/q, Keshava Nilaya, Hulimavu, Bannerghatta Road, Bangalore-560076
Hobbies and Interests	Reading Books, Learning Sanskrit, veda and Cognitive Computing

I declare that the information given above is true to the best of my knowledge and belief.

B.A DATTARAM