Project 1

Project Title	**Social and Mobile Data Integration and Analytics**
Period	10 Weeks
Location	IBM Research India New Delhi
Major	Computer Science
Mentor/Manager	LV Subramaniam
Skills Required	Java/C++ programming, mathematical maturity
Short Description	Harnessing a user's context from multiple sources plays a central role towards designing effective communication, content and services for the customer. The process needs to be PRIVACY aware to take into account the user preference from privacy as well as data sharing view points. The goal of this project is to develop analytics algorithms for scalable joins between variable-velocity mobile, social and enterprise data to gather the users context.

Project 2

Project Title	**Location and Context Aware Recommender Systems**
Period	10 Weeks
Location	IBM Research India New Delhi
Major	Computer Science
Mentor/Manager	LV Subramaniam
Skills Required	Java/C++ programming, mathematical maturity
Short Description	How can we personalize the interactions with individuals based on our awareness about their current location, current transaction and transaction history, propensity and interests? Also how can we utilize the awareness to give contextual recommendations to the user?

Project 3

Project Title	**Graphiti: Big Text and Graph Analytics**
Period	10 Weeks
Location	IBM Research India New Delhi
Major	Computer Science
Mentor/Manager	LV Subramaniam
Skills Required	Java/C++ programming, mathematical maturity
Short Description	The goal of this work is to enable machine learning, text and graph analytics @ scale. The goal is to develop a graphical understanding from global analysis of text data and then analyzing the resulting graph structures to derive higher level inferences. Today social media, telecom graphs and other large graphs have richly labelled nodes and edges resulting from the rich content that accompanies the link structures and analyzing this is important in many applications.

Project 4

Project Title	**Smart Matching of Data/Analytics to Cloud Platform**
Period	10 Weeks
Location	IBM Research India New Delhi
Major	Computer Science
Mentor/Manager	LV Subramaniam
Skills Required	Java/C++ programming, mathematical maturity
Short Description	The application and data makes certain demands on the platform. Till now this has never been addressed in an automated fashion. Our goal is to understand the characteristics of the data and the analytics to determine the models and design patterns for public clouds, private clouds, and PureApplication systems.

Project 5

Project Title	**Exploiting Social, Enterprise & Mobile data to Generate Recommendations on Big Semantic Graphs**
Period	12 Weeks
Location	IBM Research India New Delhi
Major	Computer Science, Electrical Engineering
Mentor/Manager	Sameep Mehta
Skills Required	Knowledge of any Programming Language, Data Structures and Graph Algorithms.
Short Description	In this project, we will explore the recommendation generation (link prediction) problem on Big Data (represented as graphs). We will consider multiple sources of information and fuse them to form a single semantic graph. The graph will have nodes and edge of multiple types. For example, nodes could represent people, place, skills, car type, movies and links will be association between the nodes. The goal will be to develop novel link prediction algorithms which can leverage not only the graph structure but also the meaning of nodes and edges. The different data sources which we can take into account are Social (Twitter), Enterprise Data (structured) and Mobile data (Real Time Spatio Temporal). Apart from accuracy of the algorithm, we will also focus on scalability & latency of the proposed system and would ideally port the algorithms on BigData platform.

Project 6

Project Title	**CrowdScreen**
Period	12 Weeks
Location	IBM Research India Bangalore
Major	Computer Science
Mentor/Manager	Diptikalyan Saha
Skills Required	Operating Systems, Compilers, Java
Short Description	This project will explore how to combine multiple mobile screens together into a bigger display such that any content run in any one of the mobile device can be seen seamlessly in the bigger display. The display can dynamically grow by adding more mobile phones.Interested candidates should be proficient in Android based mobile app development and experience in visualization project using libraries like OpenCV.

Project 7

Project Title	**Goal mining and variability analytics in product offerings/services**
Period	12 Weeks
Location	IBM Research India Bangalore
Major	Computer Science
Mentor/Manager	Karthikeyan Ponnalagu/ Shantanu Godbole
Skills Required	Text Mining and Analytics, Goal Model representation as Graphs (F-arc Graph for ex.), UML , OCL, Ontologies and Domain Markup Languages, Java, Eclipse Plugin, Architectural patterns.
Short Description	Perceived client organization goals or business outcomes play a central underpinning role in engineering product offerings and process services in development organizations such as IBM. There is also strong motivation from an IBM strategical stand point to develop offerings and provide services to enhance IBM's customer businesses and fulfill their growth priorities. This mandates that the business functions (also viewed as services that enable a customer or customer's customer) need to be specifically designed with high variance and subsequently flexible in terms of catering specific needs and priorities of each of our potential customers . We call this as "Goal oriented Variability Management ". In current state of the art, goal models or business motivation models (BMM, an OMG standard) are assumed to be available in a AND OR relationship based graph structures . They are subsequently employed for product Line engineering using feature models or service oriented variability modeling and management practices. But as we observe, generation of such goal models and subsequent adjustment for changing client priorities or client themselves is a costly and time consuming task . In such scenarios, with a vast porfolio of offerings and services with an organization like IBM, it become a difficult exercise to correctly identify the relevant

	offerings that enable client's business and subsequently the flexibility aspect of how best the offerings or services can accommodate the changing needs of the customer. This is purely a big ask for a business consultant or client facing team. Therefore, In the scope of this project, we would like to build a platform to focus on Goal /Business Motivation Model extraction from system specifications, User interactions, market materials, test cases and so on and generate goal/sub goal entities with AND/OR relationships. Subsequently the generated graphs will be pruned to identify the scope of variations of offerings/services that can augment the goals of the client organization and the specific variability needs and options that such offerings/services can or need to accommodate. Also such a platform, can enable how future offerings can be conceptualized and engineered specifically to address the gaps in fulfilling the prioritized business needs of a given industry or a customer organization belonging to an Industry.

Project 8

Project Title	**Spatio-Temporal Analysis of Epidemic Outbreaks**
Period	12 Weeks
Location	IBM Research India Bangalore
Major	Computer Science
Mentor/Manager	Vinayaka D Pandit/ Pankaj Dayama
Skills Required	R/SPSS
Short Description	The incidence of an epidemic is collected as spatio-temporal data, where each incidence is tagged with a unique spatial and time identifier. For a given temporal window (say weekly), one can have different spatial windows defined by administrative or political boundaries like postal codes, constituencies, districts, regions, etc. The raw incidence data, when aggregated at a spatial window, say postal code, results in multiple time series, one for each postal code. As the spatial window becomes larger, the number of time series data decreases, with a single time series for complete aggregation (for example country). Thus at one end of the spectrum of spatial windows, there is multiple time series, whose count decrease to one at the other end of the spectrum. One can expect that at finer spatial resolutions, the epidemic dynamics is governed by spatial process. The influence of the spatial process decreases as the spatial window becomes coarser, finally ending with a single time series. The analysis of influence of the spatial process will help in understanding the disease dynamics at different spatial resolutions, which in turn help in disease control mechanisms. We will model the data based on Dengue disease dynamics. The intern is required to have good theoretical and working knowledge

	of regression, causality, and time series.

Project 9

Project Title	**A compositional approach to generalize wordnet to phrase-net, concept-net and sentence-net**
Period	12 Weeks
Location	IBM Research India Bangalore
Major	Computer Science
Mentor/Manager	Vikas Raykar
Skills Required	Natural language processing, Machine learning
Short Description	WordNet is a large lexical database of English where words are grouped into sets of cognitive synonyms (synsets) and various semantic relations among words are available. The goal of this project to leverage on wordnet and automatically build (with minimal supervision) a similar resource for phrases, concepts, and sentences. Specifically the project would involve building models to predict the type of relation (equivalence, generalization, specialization, implication, support etc.) between two phrases, concepts, and sentences. Ideally a compositional approach would be preferable.

Project 10

Project Title	**Prediction Markets and Learning for Enterprise Decision Making**
Period	12 Weeks
Location	IBM Research India New Delhi or Bangalore
Major	Computer Science, Operations Research
Mentor/Manager	Dinesh Garg
Skills Required	Convex Optimization, Game Theory, Mechanism Design, Statistical Learning Theory, Online Algorithms
Short Description	There is a thin line between "Prediction" and "Forecasting". Prediction is a subjective belief of an individual regarding outcome of a future event (e.g. outcome of an election). It is typically based on an individual's knowledge, information, and wisdom. The forecasting, on the other hand, is an objective assessment about the outcome of the same future event; and it is based on the statistical nature of the data. In an enterprise, there are several complex decision making scenarios where its not only plausible but also necessary for both human and forecasting algorithms to interact with each other and improve the quality of the decision making. In recent times, prediction markets have become a popular choice for making prediction through harnessing the wisdom of society. Learning algorithms, on the other hand, are default choice to support forecasting. In this project, our goal is to understand the

	intriguing connections between prediction markets and learning algorithms. We want to understand how a cross pollination between these two disciplines can help take better decisions in select enterprise applications.

Project 11

Project Title	**Motivate Energy Conservation using Mobile-based Augmented Reality Games**
Period	12 weeks
Location	IBM Research India Bangalore
Mentor/Manager	Mohit Jain
Major	Computer science
Skills Required	Android programming, Game development experience, HCI skills
Short Description	We aim to design, develop and study mobile-based augmented reality games to motivate people about energy conservation, by making them self-aware and to self-discover ways to conserve energy. The consumer can use the phone camera to point at different energy-consuming appliances, the phone app will identify the appliance, provide energy rating, and estimate energy consumed by that appliance during a month (based on the tariff structure of that location using GPS). Moreover it can suggest alternative appliances with less energy consumption, along with the cost and breakeven point for the alternatives.

Project 12

Project Title	**Collaboration supported using Multiple Large Displays with Different Interaction Settings**
Period	12 Weeks
Location	IBM Research India Bangalore
Mentor/Manager	Mohit Jain
Major	Computer science
Skills Required	Kinect programming, Hacking skills, HCI skills
Short Description	Collaboration in meeting rooms with large displays is a common phenomenon for decision-making. In this work, we propose and develop novel (individual and group) interaction techniques to support collaboration for utility companies related decision-making tasks. We explore the interaction space offered by multiple large displays, a touch-enabled tabletop, multiple tracking devices, for multiple people collaboration. We use Microsoft Kinect for body tracking, and Leap Motion and Myo for tracking hand and finger. The interactions depend on various factors including proximity to

	the screen, proximity to other users, users switching roles (leader, non-leader, listener, helper, etc.), users switching between different interaction settings (close to screen, sitting close to touch-enabled tabletop, standing, etc.), etc.

Project 13:

Project Title	**Augmenting Physical Learning with Virtual information through google glass**
Period	12 Weeks
Location	IBM Research India New Delhi or Bangalore
Major	Computer Science, Human Computer Interaction, Wearables
Mentor/Manager	Nitendra Rajput, Danish Contractor
Skills Required	Android programming, HCI, signal processing
Short Description	What if you could build a next generation classroom where the teacher wears a google glass and through it, can engage more effectively with the student – (a) by identifying the student through the google glass and then retrieving information about the user's interesting characteristics on the glass, and thus engaging more contextually with the user, (b) by getting a nicer explanation of the user's question through retrieval of a illustrative video that can then be rendered to the student's phone or tablet or glass as an answer. If you wish to make this reality happen, come join us with a bag full of HCI techniques, some android programming and a lot of enthusiasm.

Project 14:

Project Title	**Mobile Sensor Data Analytics**
Period	12 Weeks
Location	IBM Research India New Delhi or Bangalore
Major	Computer Science, iOS/Android
Mentor/Manager	Nitendra Rajput
Skills Required	Classification, Clustering, mobile application development
Short Description	We have a large amount of user interaction data that is sourced from the various sensors on amobile device (accelerometer, gyroscope, mic, light sensor, GPS, barometer, etc.). The goal of this project is to build analytics engines that can efficiently mine this data to derive interesting insights and take those insights into making some cool apps on the mobile devices.

Project 15:

Project Title	
Period	12 Weeks
Location	IBM Research India New Delhi or Bangalore
Major	Computer Science, iOS/Android
Mentor/Manager	Malolan Chetlur
Skills Required	Classification, Clustering, mobile application development
Short Description	This IBM Research Internship is a position with the Smarter Education group in India Research Lab for two and a half month (10 weeks) to employ IBM Analytics capabilities in the area of machine learning, natural language processing and statistical analyses in the domain of online Education, particularly with MOOCs dataset.

Project 16:

Project Title	**Information Representation and Extraction from Financial Data**
Period	12 weeks
Location	IBM Research India New Delhi
Mentor/Manager	Arvind Agarwal
Major	Computer science
Skills Required	Natural Language Processing, Machine Learning, Python, Scikit-learn
Short Description	We need tools and technology to understand and process financial data. While there is a plenty of structured financial data, most of the real time financial data that has a direct impact on financial market is still unstructured, such as real time events. One of the challenging problems in assessing the current market trend is to understand the vast amount of unstructured text. We propose a project where the intern would (1) identify what an important event means in the financial domain; (2) represent and extract events using natural language processing and information extraction tools

Project 17:

Project Title	**Policy Non-compliance Alert Platform**
Period	12 weeks
Location	IBM Research India New Delhi
Mentor/Manager	Arvind Agarwal / Arun Kumar
Major	Computer science
Skills Required	Natural Language Processing, Machine Learning, Python, Scikit-learn

Short Description	Keeping up with the latest policy changes has been one big challenge for large companies. Usually, there is a huge downside involved in the case of non-compliance. Two main departments of an organization that get impacted the most are Finance and HR. These departments have to constantly keep up with the information being published in various media to make sure that they comply with the latest government policy changes. Further, where there regulations, transactions need to be constantly monitored for any violations. We propose to build a platform that would facilitate policy compliance in these scenarios.

Project 18:

Project Title	**Event Impact Expectation Prediction**
Period	12 weeks
Location	IBM Research India New Delhi
Mentor/Manager	Arvind Agarwal / Arun Kumar
Major	Computer science
Skills Required	Natural Language Processing, Machine Learning, Python, Scikit-learn
Short Description	Important events such as recent “Fed interest rate hike” have a huge impact on the financial market. The intensity of the impact depends on that expectation that market (general public, large financial entities, media) carries. Depending on the mismatch between the expectation and the outcome of the actual event could be positive or negative. Furthermore, various events may have different levels of impact on the overall economic conditions, all of which may collectively impact how the financial instruments move. In this project we would analyze the internet data to predict the sentiment/expectation of an event of interest. The project would use NLP, IE and ML tools for the prediction.

Project 19:

Project Title	**SkillConnect: A Dynamic Enterprise Expertise Taxonomy**
Period	12 Weeks
Location	IBM Research - India, New Delhi
Major	Computer Science, Data Mining
Mentor/Manager	Rakesh Pimplikar / Gyana Parija
Skills Required	Information Extraction, Classification & Clustering, Text Mining, Advanced Java
Short Description	Knowing expertise within an organization is very important to assess gaps/gluts in terms of skills required for upcoming projects. Moreover, organizing expertise and skills in a taxonomy, with possible levels being Capability – Competence – Expertise – Skills – Knowledge, enables an enterprise to take decisions on formation of right project teams, upskilling or reskilling of employees, hiring

	experts from external sources, employee compensation vis-à-vis market trends, and so on. Manual creation of such a taxonomy can provide just a snapshot of an enterprise expertise, whereas set of expertise and skills present in an organization keeps on changing every day. Hence, it makes sense to create a taxonomy dynamically, so that it always provides a current view of enterprise expertise and enables decision makers to take the right decisions. This project aims at building a solution that 1) takes several data sources, like employee profiles, projects, publications, accomplishments, etc., as input, 2) infers capability, competence, expertise and skills for every employee, and 3) organizes them in an expertise taxonomy.