Research Statement

I am an experimental computer scientist, who likes to develop novel and practical human behavioral event sensing technologies that, (1) capture observable low-level physical signals from our bodies and surrounding environments in the form of acoustic and electromagnetic waves, (2) map them to relevant physical and behavioral measurements, and (3) employ a deep understanding of new deep learning, machine learning, signal processing and natural language processing techniques (on transcribed speech-text), to rethink the core mechanisms of existing human behavioral event sensing technologies.

My primary area of expertise is **Developing Sensing Processing Technologies** which overlaps with **Applied Deep Learning and Machine Learning, Human Centered Computing (**e.g. Health Computing, Human Machine Interaction, and Wellness Monitoring Applications**), Internet of Things (IoT), Cyber-Physical Systems (CPS), and Natural Language Processing**. I enjoy building **data-driven**, **application-specific novel technologies,** as well as **new systems and applications,** that involve sensors, mobile devices, and cloud services. Research challenges that I deal with are the uncertainties in physical world sensing, human factors such as the user-context and mobility, limitation of current technologies, and resource constraints of the sensing data and platform. Technologies and Systems that I have developed are **human-centric**, several of them are attributed to **health and wellness**, and in general, they are in the scope of **Ubiquitous Computing**.

Research Philosophy

My research philosophy is built upon four main principles: *novelty*, *practicality*, *impact, and multidisciplinary*.

- **Novelty:** I aim at the 'impossible' while trying to envision technologies that are useful but do not yet exist. Advancement in technology, such as a new learning algorithm, or a new lower bound of a useful algorithm, a new sensing device help me predict *what's the next exciting thing* that I could build with my expertise.
- **Practicality:** I put much of my effort towards studying the problem domain in order to design and implement a real and working approach (technology), rather than just a research prototype. While the vast body of work in machine learning, deep learning or NLP have achieved high accuracy in controlled experiments and datasets, readily applying these works in realistic settings yields poor results. A big part of my research is to identify the challenges in realistic systems (applications), trying to understand the physical nature of it and design an approach or technique that is capable of identifying the variability inside and outside the system, adaptive to the uncertainties and constraints of both human and systems.
- **Impact:** While deciding on what research idea to invest time on, I put much weight on its impact. I get excited to work on problems whose end results have direct effects on people's lives. Novel techniques that potentially have wide spread usage and help improve a great many people's lives are of special interest to me.
- **Multidisciplinary:** While computer science is interdisciplinary in nature, today's most salient areas of technical research, such as, Data Analytics, Internet of Things - are multidisciplinary. A multidisciplinary group brings together the expertise and novel perspectives from different fields that can lead to creative and high impact research. I have the experience of working with collaborators from the fields of Computer Science, Psychology, Behavioral Science, Mechanical Engineering and Electrical Engineering. These collaborations have enabled me to address research challenges from novel eccentric ways, establishing a 'common language' with people from different disciplines and learning from scientists whom I would never previously had a chance to interact with.

Recent Research

Detecting Mental Disorders: Mental health problems, such as depression and social anxiety disorder, are often under-diagnosed and under-treated, in part due to difficulties identifying and accessing individuals in need of services. One challenge is, requiring individuals to self-report their symptoms, which depressed and anxious individuals may not be motivated to do. These detection methods are also vulnerable to social desirability and other subjective biases. Identifying objective, non-burdensome markers of depression and social anxiety, such as features of speech, could help advance assessment, prevention, and treatment approaches. Prior research examining speech detection methods has focused on fully supervised learning approaches employing strongly labeled data. However, strong labeling of persons high in symptoms or state affect in speech audio data is impractical, in part because it is not possible to identify with high confidence which regions of a long speech indicate the person's symptoms or critical affective state. I developed a **weakly supervised deep learning framework** for detecting social anxiety and depression from long audio clips. Specifically, I presented a novel **feature modeling (knowledge engineering and representation)** technique named **NN2Vec**, which identifies and exploits the inherent relationship between speakers' vocal states and symptoms/affective states. In addition, I developed a new **multiple instance learning adaptation of a BLSTM classifier,** named **BLSTM-MIL**. My novel framework of using NN2Vec features, with the BLSTM-MIL classifier achieves significantly higher F-1 scores in detecting speakers, who are high in social anxiety and depression symptoms. This work has been published in the flagship conference of ubiquitous computing: **Ubicomp and IMWUT 2018**. This work is the core part of the **NSF Smart and Connected Health Grant** (2018) in collaboration with *Nursing and Behavioral Science* department of *University of Tennessee*.

Distant Emotion Recognition (DER): Distant emotion recognition (DER) extends the application of speech emotion recognition to the very challenging situation, that is determined by the variable, speaker to microphone distance. The performance of conventional emotion recognition systems degrades dramatically, as soon as the microphone is moved away from the mouth of the speaker. This is due to a broad variety of effects, such as, background noise, feature distortion with distance, overlapping speech from other speakers, and reverberation. I developed a novel solution for DER, addressing the key challenges by identification and deletion of features from consideration which are significantly distorted by distance, creating a novel **feature modeling (knowledge engineering and representation)** technique, called **Emo2vec** and overlapping speech filtering technique, and the use of an **LSTM** (deep learning) classifier to capture the temporal dynamics of speech states found in emotions. A comprehensive evaluation is conducted on two acted datasets (with artificially generated distance effect), as well as on a new emotional dataset of 12 spontaneous family discussions, with audio recorded from multiple microphones placed in different distances. This work has been published in the flagship conference of ubiquitous computing: **Ubicomp and IMWUT 2017**. In near future, the work will be integrated in a monitoring system for family eating dynamics research in collaboration with department of *Behavioral Science* in *University of Southern California*.

Detecting Agitated Vocal Events (DAVE): I developed a system that continuously monitors and detects agitated vocal events, which is useful for the elderly population suffering from Dementia. DAVE, using a novel combination of **acoustic signal processing** and **multiple text mining** techniques, automatically detects and records the 8 major vocal agitations for dementia patients, as defined by the medical community. This includes cursing or verbal aggression, constant unwarranted request for attention or help, negativism, making verbal sexual advances, crying, screaming, laughing, and talking with repetitive sentences. The novelty of DAVE includes the comprehensiveness of addressing all 8 vocal events, using the text of the vocalizations only when accurate, combining text and acoustic features when necessary, and employing text mining and feature identification. Additionally, to understand the ambiguity of spoken words in natural language, I developed a novel **word sense disambiguation** technique, **adapting the Lesk algorithm**. Unlike many other systems, it does not need any wearable or in-situ sensors (other than Kinect sensor), thus it should not cause discomfort for patients. The scope of this project can be extended to smart homes and security, where DAVE can be used to detect several previously undetectable verbal anomalous events. This work has been published in the new and leading conference of connected health: **CHASE 2017** and served as a key element in obtaining an **NSF Smart and Connected Health Grant.**

Sensor Data Integration: Conventional multimodality sensor data integration approaches concatenate features from different sensory streams (from a detection window) and feed the concatenated feature-set to a supervised learning classifier. Since different sensors perceive only a sub-part of an event, a significant part of the sensory signals (from each devices) in an event detection window, contain irrelevant (noisy) information, which makes supervised learning classifiers perform poorly. I have developed a novel and general **multidimensional constrained multiple instance learning neural network** approach to integrate information at the decision-level from multi-modality multi-point wearable sensing data and applied on automated human activity detection task. This approach considers each of the sensory streams (signals from sensor devices) in an activity detection window as weakly labeled and extracts knowledge through an integrated weakly supervised learning approach from these combined streams. Conventional multiple instance learning (a form of weakly supervised learning) assumption does not entirely hold up to the characteristics of the sensory streams in human activity event detection windows. Hence, we developed a novel **Constrained Multiple Instance Learning** approach, incorporating the attributes of sensory stream data to perform weakly supervised learning for the targeted task. The work is currently under review at the flagship conference of ubiquitous computing.

AsthmaGuide: There has been an increased use of wireless sensor networks in the medical sector. I developed a system, named AsthmaGuide, in which a smart phone is used as a hub for collecting physiological, environmental, human input, picture, and video information from several wireless sensors like sensordrone, electronic stethoscope, pulse oximeter, etc. The data, including data over time, is then displayed and analyzed in a cloud web application for both patients and health-care providers to view. AsthmaGuide also provides an advice and alarm infrastructure based on the collected data and parameters set by the health-care providers. We are currently in the process of deploying the system in a *Children's hospital* in *South Korea*. This work has been published in the leading conference of wireless, connected and mobile health research: **Wireless Health 2016 (nominated for best paper)**.

Early CKD detection: In this study I have considered 24 predictive parameters and have developed a machine learning classifier to detect Chronic Kidney Disease (CKD). The approach had achieved 69% reduction of mean square error compare to the state of the art (CKD-EPI equation) GFR estimator. I also have performed feature selection to determine the most relevant attributes for detecting CKD and have ranked them according to their predictability. This work has been published in the leading conference of healthcare informatics: **ICHI 2016**.

Recognition and Impact

My research has been well-accepted by a diverse set of communities. I have regularly published my works in top-notch conferences in the fields of Ubiquitous Computing, Connected Health, and Wireless Sensor Networks. Recently, I have received the **Graduate Student Award for Outstanding Research** from UVA Department of Computer Science. In 2016, my work, AsthmaGuide, was **nominated for the best paper award** in the Wireless Health 2016 conference.

I have gained the experience of working in two leading research labs (**Nokia Bell Labs, Bosch Research**) as a research intern. During my *research internship* in the *`Human-Machine Interaction'* group at *Bosch Research (Palo Alto, CA)*, I developed a **Robust Audio Event Detection System**, a hierarchical approach that detects the presence of an event in an audio clip using a *Dilated Convolution Neural Network* and identifies the event boundaries via a novel feature modeling technique (*knowledge engineering and representation*) and *bi-directional many to many LSTM* classifier. Currently, this work is submitted in a leading conference. Additionally, an **US patent** has been submitted for this work. During my *research internship* with the *BHAG Realization Lab at Nokia Bell Labs (Murray Hill, NJ, USA)*, I developed natural human-machine interaction techniques (using **speech analytics, NLP and IMU sensor analysis**) for **distributed cyber physical systems** (including **drones, robots**, etc.).

Several undergraduate and graduate students have used the outcome of my research to find their own research directions. My work on AsthmaGuide has helped one graduate (masters) and one undergraduate student to complete their theses. My work on Distant Emotion Recognition (DER) is another remarkable project which is currently being used by at least three students, who are researching on how to use the system in solving a variety of problems, such as, distant anxiety detection, multi-person in home emotion detection, etc.

Based on two of my works, one on agitated vocal event detection (DAVE) and the other one on detecting mental disorder from weakly labeled data, we have so far been awarded two research grants amounting *$1,200,000* and *$180,000* from **NSF** (Award Number: 1838615) and a **funding agency in Korea**, respectively. I have actively contributed in conceptualizing and writing the grants, which I believe have greatly enriched my experience to be a successful researcher.

Future Agenda

I envision a future where technology will become invisible yet omnipresent. In that world, smart dust will be a reality, every object will have tiny embedded computers that are barely visible, and human-device interactions will be via speech, gestures and brain-computer interfaces. With the determination to realize this vision by making significant contribution to the field, I would like to address the following related but separate thrusts in near future:

Internet of Things (IoT)

With the increasing interest in smart cars, smart homes, smart health-care, and smart cities, we are moving toward a smart world. Companies are developing devices, such as, smart glasses, smart watches, smart rings, smart earphones, and smart forks and plates, etc. In near future, these ultra-low power electronic gadgets will serve applications ranging from entertainment to productivity to wellness monitoring, and these smart gadgets will talk to the wearable computers using wireless channels. The current trajectory of the numbers of smart devices being deployed implies that eventually trillions of things will be on the Internet. In an Internet of Things (IoT) world, there will exist a vast amount of raw data being continuously collected. In practice, an IoT System receives noisy, delayed, and incomplete information on the system state. It will be necessary to develop techniques that convert this raw data (containing noise, ambiguity and partial aspects of sensing) into usable knowledge. The main challenges for data interpretation and the formation of knowledge include addressing noisy, physical world data; adaptive knowledge engineering and representation; effective data integration; and developing new scalable inference techniques that do not suffer the limitations of Bayesian or Dempster-Shafer schemes. These limitations include the need to know *a priori* probabilities and the cost of computations. In near future, a whole new set of research problems will emerge, such as, how to coordinate a wide number and variety of smart devices producing wide range of multimodality data; how to develop distributed sensing gadgets and technologies, and effectively combine information on-the-fly; what new applications can be created; and how to develop applications that run seamlessly in all the devices. I want to leverage my expertise in machine learning, deep learning, signal processing and data-integration to address these exciting problems.

Modeling Human Behavioral Events

Internet-of-Things (IoT) has created a wave of opportunities for modeling human behavioral events. Supported by rapid innovations in machine learning, signal processing, computing, and wearable-systems, the concept of connected sensing is redesigning almost every aspect of our lives. Innovating novel, low cost and noninvasive sensing techniques to model/identify individual’s behavioral state (i.e., emotions, mental disorder, etc.) has become one of the core research areas. Accurate human behavioral state detection can facilitate assessment of stress or competence of factory workers, caregivers in nursing homes or drivers in smart-cars and avert hazardous situations. Moreover, detection of mental states can improve assessment and treatment of mental disorders, monitoring and care of patients suffering from agitation, dementia or stroke rehabilitation. Yet, accuracy of detection or modeling human behavioral events is far from desirable to have any practical implication. This is due to the lack of adaptability of the current state-of-the-art machine learning, data analytics techniques, with the characteristics and constraints of human behaviors. The core of my work lies in building everyday sensing and prediction techniques that are adaptive to the uncertainties and constraints of human and system. Currently, I am working with the department of Physiology of UVA to understand the semantic contents of mental-disorder patient’s speech from limited training data that highly indicate their respective mental-disorders. The findings of this research can be used as an objective biomarker of mental disorder symptoms. To continue my quest of developing human behavioral event modeling techniques with high enough accuracy for practical implementation, I aim to push the boundary of current machine learning, deep learning, data analytics techniques with a deeper understanding of human behavioral characteristics and constrains of non-invasive human sensing systems. I believe my experience in modeling behavioral events (i.e., emotion, depression, social anxiety, etc.), machine learning, deep learning, sensor data analytics and NLP will help me in my quest.

Humans and Smart Cities

Over half of the human population live in cities today and this number is estimated to increase to 70% by 2050. With the new technological trends that are emerging, such as, Big Data, Internet of Things and Ubiquitous Communications, all leading to a digital society, new ways of living and sharing knowledge are occurring. Using Internet of Things (IoT), smart cities are sensing what is happening in a city - parked cars, traffic jams, hospital beds available, energy consumption, water or air quality, temperature, noise, etc. In current implementation and practice, the smart city vision exploits a very technology focused approach with the understanding of the city as a machine and allows for acting in the real word as to adapt it to new circumstances. The potential of sensing and integrating human activities in the equation is yet to be explored. An effective smart city needs to sense humans in personal and social level and adapt the information in its decision making. Hence, there is a need of novel scalable integration of supervised sensing from multimodality multidimensional data with reinforcement learning based decision-making approaches. An example can be an extension of my previous project AsthmaGuide. A personalized smart city application can allow the patients with respiratory problems, such as, asthma, allergies, etc., to select the best route (with the least possible impact to their medical problem) for their commuting between two places. A personal mobile sensing technology would assess the severity of the patient's health, get the real-time measurements of pollution, pollen, etc. as well as traffic congestion, public events, etc. from the city sensors (via servers) which are close to the potential routes, and calculate the most feasible ones.

If properly understood and used, this huge amount of human and social sensing data has the potential to bring enormous value and important improvements to cross-concerning areas that have strong and direct impact on the quality of people's lives, such as healthcare, urban mobility, public decision making and energy management, etc. Consequently, a new set of research challenges will emerge, such as, effective and noise adaptive knowledge engineering/feature modeling to extract meaningful data from millions of dimensions, effective integration of discrete and continuous data from millions of sensing dimensions with sampling rate ranging from once a day (measuring blood pressure of a patient) to 44 kHz (video), non-invasive human-in-the-loop feedback technologies for smart governance, only to mention a few. I believe, my experience in human behavioral event sensing from multidimensional sensing data, mobile-health sensing from projects like AsthmaGuide or DAVE and developing adaptive feature modeling/knowledge engineering from DER or mental disorder detection projects will help me in addressing these research challenges.

Big Data Analytics

Recently, Big Data has become one of the most active research areas. As researchers in sensing and ubiquitous computing, we are a major producer of data, which are generated by long running sensing systems and/or contributed by millions of devices as in crowd-sensing applications. These unstructured continuous data come in all different modalities (text, audio, electromagnetic signals, etc.), and the significant portion of these unstructured multimodal sensing data contains noise and ambiguity. Since, state-of-the-art machine learning and deep learning techniques are not adaptive of having large amount of noise and ambiguity in the training and testing data, a significant effort in big data analytics is spent on cleaning the data. An interest of mine is therefore to apply my experience in multimodal data analytics (audio, text, IMU sensors, time-series data, etc.) in conjunction with experience in research on decision-level multimodal noisy data integration technique, to develop adaptive and scalable machine learning approaches, that can extract meaningful knowledge from noisy, ambiguous big unstructured data.