Transcript of the Medical AI Presentation by
Dean Jiajie Zhang, PhD

Slide 1:
Hi, everyone. I am Jiajie Zhang, Dean of the School of Biomedical Informatics at The University of Texas Health Science Center at Houston. Our school is one of the largest academic programs of biomedical informatics in the nation and the only one as a free standing school. We are located in the Texas Medical Center, the largest healthcare cluster in the world. At our school, we do education, research, and application in Medical AI and Data Science, Clinical and Health Informatics, and Bioinformatics and Systems Medicine.

Today, I talk about Artificial Intelligence applications in medicine. I want to tell you why **AI is to Medicine Today What the X-ray was to Medicine a Century Ago, and Much More…**

Slide 2:
Here are the sub-topics that I will go through:

- Medical AI is the X-ray of the 21st Century.
- Medical AI is real, finally.
- Medical AI is easy.
- And Medical AI is hard.
- Finally, I will explain what the success of Medical AI requires deep clinical integration.

Slide 3:
Medical AI is the X-ray of the 21st Century.

Slide 4:
A century ago, X-ray enabled doctors to see invisible structures inside the body.

Slide 5:
Today, AI is enabling doctors to not only see, but predict, previously unidentified patterns within massive medical and biological data. This is a futuristic picture of what medicine in the hi-tech world looks like. The key physiologic states of the body are patterns of data that intuitively visible and easily actionable, with the touch of a finger. The data are dynamic, can go down to molecular levels and can go up to population levels. The data can also go back in medical history and go forward as predictions of health status of the patient. Diagnosis can be made and treatment plans can be generated, with the help of artificial intelligence.

Slide 6:
Artificial intelligence, since its beginning in the 1950s, promised a lot of things. But nothing really worked, including expert systems for medicine, which is a big part of early AI. Today, three drivers made AI a reality and capable of solving real world problems, sometimes better and faster than people. These three triggers are massive, massive data in electronic and computable forms, powerful and affordable computing, and universal connectivity.

Slide 7:
Let us take a look at what happened over the past 5 to 10 years. The amount of data are increasing exponentially; for medicine, adoption of electronic health records, or EHR, is almost universal. The cost of sequencing the entire human genome dropped from the cost of a Boeing 737 to the cost of a smartphone. Of course the speed to internet connection and the speed of computing are both massively increased. Whether you talk about Industry 4.0 or Health 2.0, we are at a historically unique moment – the cognitive revolution that is liberating people from the cognitive labor. This is as fundamental as the industry revolution that started more than 200 years ago and liberated people from physical labor; and the agricultural revolution more than 10 thousand years ago that provided people with food security through the transformation from hunting and food gathering to farming.

Slide 8:
US Senator Ben Sasse summarized the moment we are in today very nicely in an article in the Wall Street Journal: “The past 20 or 30 years, and the next 20 or 30 years — really is historically unique. It is arguably the **largest economic disruption in recorded human history.**” Computing technology has transformed or has been transforming all major industries, from information access, communication to retail, entertainment, to travel, finance, and to knowledge intensive education.

Today, it is finally getting into the healthcare industry, fundamentally transforming how we treat patients, how we make diagnoses, how we prevent diseases, and how we make biomedical discoveries.

Slide 9:
The list of medical AI applications is expanding rapidly. Imaging segmentation, imaging annotation, and imaging based diagnosis are becoming the early playground of AI applications. Prediction is one other major strength of machine learning. Examples include the prediction of sepsis onset before it is detected by regular algorithms, predicting disease progression, calculating risks for various medical conditions such as myocardial infection, and heart failure, discover computational biomarkers to detect medical conditions such as Parkinson’s from typing on a keyboard. Of course natural language processing is a very active application area of AI. Medical AI systems can even take the medical licensing exam, performing better than 80% or more of human takers. The list goes on and on and is potentially unlimited. Let’s take a look at a few examples of medical AI applications that are developed by the researchers at our School of Biomedical Informatics at the University of Texas Health Science center at Houston.

Slide 10:
Developing a new drug typically takes more than 10 years and costs more than a billion dollars. If an approved drug on the market for one disease has functions that can be potentially used to treat a different disease, that is, repurposing an existing drug, it will benefit the patients and the medical community. Medical AI can help with discovering the signal for a new disease through data mining of massive medical records.
Dr. Hua Xu, through the use of natural language processing and data mining over millions of patient records, discovered that Metformin, which is one a front line drug for type 2 diabetes, has a potential impact on cancer treatment. He compared cancer patients who are either diabetic or not diabetic, whether they use metformin or other diabetic drugs, and discover that diabetic cancer patients who took metformin for diabetes had a much better five year survival rate than diabetic cander patients who took other diabetic drugs and even cancer patients who are not diabetic.

This type of discovery is only possible with the availability of massive patient data and the use advanced AI tools.

Slide 11:
If you have the medical records of 50 million people for over ten years, what can you do with it? A lot! Dr. Degui Zhi used this type of dataset from Cerner Corporation to predict the risk of heart failure onset. Heart failure is a medical condition where the heart can't pump enough blood to meet the body's needs and eventually lead to death. In 2016, there are 5 million heart failure patients in US and the healthcare cost for them was $30 billion dollars. Early prediction and early prevention and treatment is the key for taking care of these patients. Dr. Zhi and his team used a deep recurrent neural network to develop a predictive model that performed really well, with an AUC around 79 to 85%.

Slide 12:
Dr. Luca Giancardo developed a machine learning model to identify computational biomarkers from the movement of hands in typing on the keyboard and touching on a smart device. He was able to use the signal to identify whether a patient has Parkinson's disease or not. This is a potentially very useful tool to do early detection and for tracking of progression of Parkinson's disease.

Slide 13:
Sepsis is the bacteria infection of the blood stream.
It is the leading cause of death in U.S. hospitals. 1 patient dies every 2 minutes in the US—more than breast cancer, prostate cancer and HIV **combined**. The good news is that 80% sepsis deaths preventable. Mortality increases 8% for every hour that treatment is delayed. So early detection of the onset of sepsis, or better yet, early prediction well before it can be detected by the current standard of care algorithms, is going to save a lot of lives. Dr. Xiaoqian Jiang, a machine learning researcher, collaborated with Dr. Bela Patel, an ICU physician, and Dr. Robert Murphy, an ER Physician, developed a machine learning algorithm based on deep LSTM that can predict the onset of server sepsis 4 hours ahead of the time. The AUC of the model achieved 92%, better than the status quo of 85% achieved by other models. Their model is currently being validated for potential deployment at our teaching hospital, Memorial Hermann Hospital at the Texas Medical Center.

Slide 14:
The "Eyes Are the Windows of the Soul," as we often hear. For medical AI researchers such as Dr. Luca Giancardo and his physician collaborator Dr. Sunil Sheth, the "Eyes Are the Windows of Health."
Retina images are typically used to diagnose eye diseases. However, they are actively developing an AI system that can use the retina image to detect signs of stroke. CT imaging is the standard of practice for stroke diagnosis. In comparison with CT machines which are very expensive and not mobile, the device for taking retina images is very small, portable, and inexpensive. The AI technology of using retina images to diagnosis stroke and other medical conditions has a lot of potential.

Slide 15:
With the development and refinement of many high level machine learning packages such TensorFlow, Trax, and PyTorch, developing medical AI applications is becoming incredibly fast and easy. With unlimited coffee and food, students, both undergraduate and graduate, can develop machine learning algorithms that can solve real world medical problems within 24 hours. Let me show you a few Datathons that our school has hosted under the leadership of Drs. Xiaoqian Jiang, Yejin Kim, and Shayan Shams.

Slide 16:
In this Hackathon in 2019, the students were given the task of predicting Sudden Unexpected Death in Epilepsy, or SUDEP. For epilepsy patients, after a seizure, there is a change of dying and this is indicated by the onset of slow activity of their EEG signals. The students were given the EEG data of real patients and asked to build a machine learning model to predict the onset of the slow wave. The best model from the students achieved an AUC of 84%, which is quite good. The student projects from the Hackathon led to five publications in a BMC special issue. Another amazing accomplishment.

Slide 17:
Just like many people in the nation and in the world, when the COVID pandemic hit, we tried to help to deal with the pandemic with our expertise. We organized a student COVID-19 Datathon. The task is to Predict COVID-19 hospitalization and mortality in Houston Metro Area.
The students were given data about

- The historical hospitalization and mortality rates
- The infection, recovery, active, and test cases (9 counties)
- The population mobility, demographics, and mask usage
- The Best model's performance achieved a Mean Squared Logarithmic Error of 16.5

Slide 18:
In the most recent Datathon we hosted last month, the students were asked to build a machine learning model to predict how well a stroke patient recovers over time.
Specifically, the students need to develop algorithms to predict changes in cognitive and Functional Independence Measure scores. Again the models the students developed performed well.
Solving a real world medical problem with real patient data within 24 hours; by undergraduate and graduate students, some of whom only learned the basics from an intro to machine learning course. This is exciting; but think about it, it is also kind of scaring. Having a model that makes good predictions, however, does not mean that makes an impact on the patients.

Slide 19:
Medical AI is still very hard.

Slide 20:
Most academic (and industrial) medical AI products never get deployed. There are many reasons why. After research and development, and after validation, both internally and externally, a medical AI product is ready for testing in the real clinical setting. However, getting the product into the production environment is one challenge. How to get it into the clinician's workflow is another, bigger challenge. What are the clinical and business utilities of the product? How to get it go through the regulatory process both internally and externally, such as FDA? How to keep the product updated with new data, modified with new algorithms? What about patient safety concerns? The list goes on and on. Let us look at a real use case to understand the complexity of getting medical AI product into the clinical setting.

Slide 21:
This case is about CT imaging for stroke.

Slide 22:
There are two types of stroke: ischemic stroke that is the blockage of the blood vessel and the hemorrhagic stroke that is the breakable of the blood vessel. Ischemic stroke accounts for 87% of stoke and there is a therapy that can save a patient's life. This therapy is called endovascular stroke therapy,

or EST, which is to insert a stent retriever from the groin all the way to the brain to remove the blood clot. CT Perfusion, or CTP, is the definitive guide to determine whether a patient is eligible for this procedure. CTP is an advanced procedure that is not widely available at smaller facilities. A relative simpler technique called CT Angiogram, or CTA, is more widely available but it is typically not good enough to determine EST eligibility. Dr. Giancardo, in collaboration with neurologists Dr. Sheth and Dr. Sean Savitz, developed a machine learning model that can use CT Angiogram to do what CT Perfusions' job.

Slide 23:
They successfully developed a CT Angiogram based AI system called DeepSymNet, which has a very good performance level. This system reduced the time from patient's ER arrival to treatment decision dramatically because the AI system only takes 1 minute to read the CT and send the report to the physicians for treatment decision.

Slide 24:
The system generates an alert after image analysis by the AI system and sent it to the stroke team by email that has an EST eligibility score along with the CT images.
The development and implementation of this system recorded a very fast "bench-to-bedside" time. It took a year to go from idea to implementation in the hospital and this is very fast.
The pipeline for this system is now integrated in 4 hospitals and it is functioning in near real time. As we can see, this AI pipeline sends alert by email which is outside of the EHR platform. Getting this system integrated into the EHR platform and the workflow of the clinicians will be next step, which is non-trivial and may well be the major barrier to its wide adoption.

Slide 25:
A technically capable AI product needs to be deeply integrated into the clinical environment and it needs to show it's clinical, operational, and financial utilities before it can be widely adopted to generate meaningful impact on patient care. This deep integration requires a lot more than AI algorithms. It requires an entire discipline called biomedical informatics; and it requires an insider's job, that is, researchers and clinicians in healthcare institution.

Slide 26:
Biomedical Informatics studies the acquisition, storage, communication, processing, integration, analysis, mining, retrieval, interpretation, and presentation of data and determines how to transform data, which are meaningless symbols, to information, which are interpreted data, to knowledge, which are validated information, and to intelligence, which are actionable knowledge, with the aim of solving problems in disease prevention, healthcare delivery, and biomedical discovery.
Biomedical informatics covers the entire spectrum of biological scales—from small molecules, genes, proteins, and cells, to tissues and organs, to individuals and populations. Biomedical informatics is a highly interdisciplinary field focused on collaborations with partners in clinical practice (e.g., medicine, nursing, dentistry, pharmacy); the biomedical sciences; public and community health; computer science and engineering; mathematics and statistics; cognitive science; social and behavioral sciences; healthcare management; and health IT policy and law. ***Data Science*** for medicine is a subfield of Biomedical Informatics; it focuses on all aspects of data for disease prevention, healthcare delivery, and biomedical discovery. ***Medical AI*** *is in turn* a subfield of Health Data Science, and it focuses on machine learning, pattern recognition, computational phenotyping, and predictive modeling.
Our School of Biomedical Informatics at the University of Texas Health Science Center at Houston is the only free-standing school among 70 or so such programs in the nation, and it is one of the largest

internationally.

The AI Revolution promises to be an exciting era. With virtually unlimited potential, medical AI is rapidly evolving to produce ever greater numbers of increasingly advanced clinical applications that will dramatically improve patient care, disease prevention, and biomedical discovery.
We are proud to be a leader in medical AI. Health data science, and biomedical informatics. ***It's great to be part of that transformation!***

Slide 27:
Thank you for your time. Hope you enjoyed this presentation. If you want to learn more about medical AI, health data science, biomedical informatics, or the education programs at our school, you can reach me by email or follow me in social media. Thank you.