UDACITY
FOR ENTERPRISE

THE SCHOOL OF ARTIFICIAL INTELLIGENCE

AI for Healthcare

NANODEGREE SYLLABUS

Overview

Power the clinical treatments of the future

Play a critical role in enhancing clinical decision-making with machine learning to build the treatments of the future. Learn to build, evaluate, and integrate predictive models that have the power to transform patient outcomes. Begin by classifying and segmenting 2D and 3D medical images to augment diagnosis and then move on to modeling patient outcomes with electronic health records to optimize clinical trial testing decisions. Finally, build an algorithm that uses data collected from wearable devices to estimate the wearer's pulse rate in the presence of motion.

A graduate of this program will be able to:

- Recommend appropriate imaging modalities for common clinical applications of 2D medical imaging.
- Perform exploratory data analysis (EDA) on 2D medical imaging data to inform model training and explain model performance.
- Establish the appropriate 'ground truth' methodologies for training algorithms to label medical images.
- Extract images from a DICOM dataset.
- Train common CNN architectures to classify 2D medical images.
- Translate outputs of medical imaging models for use by a clinician.
- Plan necessary validations to prepare a medical imaging model for regulatory approval.

Program Information

TIME
4 months
Study 10 hours/week

LEVEL
Specialist

PREREQUISITES
Intermediate Python and Machine Learning experience.

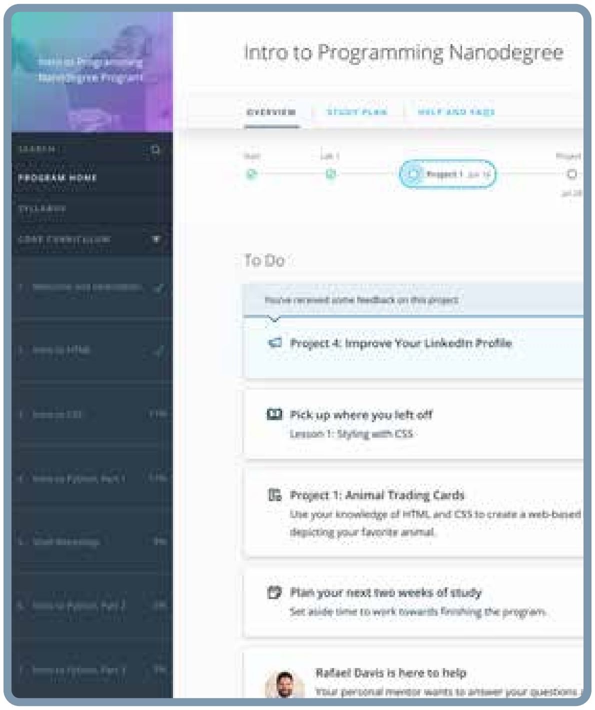
HARDWARE/SOFTWARE REQUIRED
Computer running recent versions of Windows, Mac OS X, or Linux and an unmetered broadband Internet connection.
You will use Python, PyTorch, TensorFlow, and Aequitas in this Nanodegree program.

LEARN MORE ABOUT THIS NANODEGREE
Contact us at enterpriseNDs@udacity.com.

Overview

- Detect major clinical abnormalities in a DICOM dataset.
- Train machine learning models for classification tasks using real-world 3D medical imaging data.
- Integrate models into a clinician's workflow and troubleshoot deployments.
- Build machine learning models in a manner that is compliant with U.S. healthcare data security and privacy standards.
- Use the Tensorflow Dataset API to scalably extract, transform, and load datasets that are aggregated at the line, encounter, and longitudinal (patient) data levels.
- Analyze EHR datasets to check for common issues (data leakage, statistical properties, missing values, high cardinality) by performing exploratory data analysis with Tensorflow Data Analysis and Validation library.
- Create categorical features from Key Industry Code Sets (ICD, CPT, NDC) and reduce dimensionality for high cardinality features.
- Use Tensorflow feature columns on both continuous and categorical input features to create derived features (bucketing, cross-features, embeddings).
- Use Shapley values to select features for a model and identify the marginal contribution for each selected feature.
- Analyze and determine biases for a model for key demographic groups.
- Use the Tensorflow Probability library to train a model that provides uncertainty range predictions in order to allow for risk adjustment/prioritization and triaging of predictions.
- Preprocess data (eliminate "noise") collected by IMU, PPG, and ECG sensors based on mechanical, physiology and environmental effects on the signal.
- Create an activity classification algorithm using signal processing and machine learning techniques.
- Detect QRS complexes using one-dimensional time series processing techniques.
- Evaluate algorithm performance without ground truth labels.
- Generate a pulse rate algorithm that combines information from the PPG and IMU sensor streams.

Our Classroom Experience



REAL-WORLD PROJECTS

Learners build new skills through industry-relevant projects and receive personalized feedback from our network of 900+ project reviewers. Our simple user interface makes it easy to submit projects as often as needed and receive unlimited feedback.

KNOWLEDGE

Answers to most questions can be found with Knowledge, our proprietary wiki. Learners can search questions asked by others and discover in real-time how to solve challenges.

LEARNER HUB

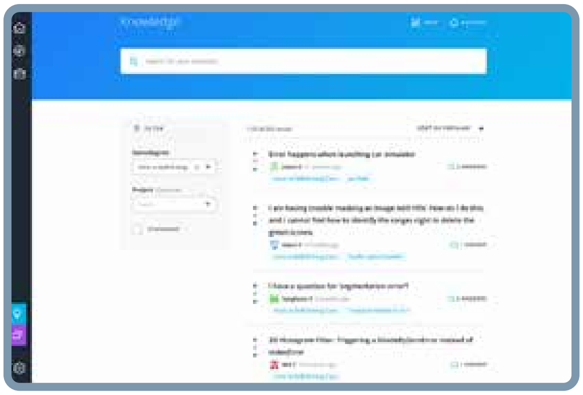
Learners leverage the power of community through a simple, yet powerful chat interface built within the classroom. Learner Hub connects learners with their technical mentor and fellow learners.

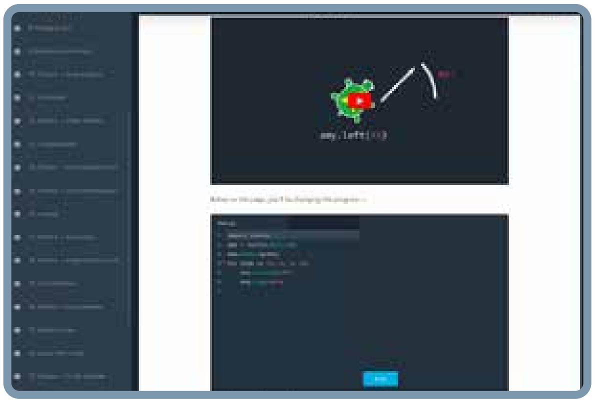
WORKSPACES

Learners can check the output and quality of their code by testing it on interactive workspaces that are integrated into the classroom.

QUIZZES

Understanding concepts learned during lessons is made simple with auto-graded quizzes. Learners can easily go back and brush up on concepts at anytime during the course.





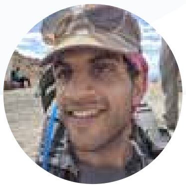
CUSTOM STUDY PLANS

Mentors create a custom study plan tailored to learners' needs. This plan keeps track of progress toward learner goals.

PROGRESS TRACKER

Personalized milestone reminders help learners stay on track and focused as they work to complete their Nanodegree program.

Learn with the Best



Nikhil Bikhchandani

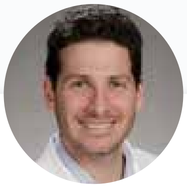
DATA SCIENTIST AT VERILY LIFE SCIENCES

Nikhil Bikhchandani spent five years working with wearable devices at Google and Verily Life Sciences. His work with wearables spans many healthcare domains including cardiovascular disease, neurodegenerative diseases, and diabetes. Before Alphabet, he earned a B.S. and M.S. in Electrical and Computer Engineering at Carnegie Mellon University.

Emily Lindemer

DIRECTOR OF DATA SCIENCE & ANALYTICS AT WELLFRAME

Emily is an expert in AI for both medical imaging and translational digital healthcare. She holds a PhD from Harvard-MIT's Health Sciences & Technology division and founded her own digital health company in the opioid space. She now runs the data science division of a digital healthcare company in Boston called Wellframe.



Mazen Zawaideh

RADIOLOGIST & CONSULTANT AT MICROSOFT RESEARCH

Mazen Zawaideh is a Neuroradiology Fellow at the University of Washington, where he focuses on advanced diagnostic imaging and minimally invasive therapeutics. He also serves as a Radiology Consultant for Microsoft Research for AI applications in oncologic imaging.

Ivan Tarapov

SR. PROGRAM MANAGER AT MICROSOFT AI

At Microsoft Research, Ivan works on robust auto-segmentation algorithms for MRI and CT images. He has worked with Physio-Control, Stryker, Medtronic, and Abbott, Ivan has helped develop external and internal cardiac defibrillators, insulin pumps, telemedicine, and medical imaging systems.

Michael DAndrea

PRINCIPLE DATA SCIENTIST AT GENENTECH

Michael is on the Pharma Development Informatics team at Genentech (part of the Roche Group), where he works on improving clinical trials and developing safer, personalized treatments with clinical and EHR data. Previously, he was a Lead Data Scientist on the AI team at McKesson's Change Healthcare.

Nanodegree Program Overview

Course 1: Applying AI to 2D Medical Imaging Data

2D imaging, such as X-ray, is widely used when making critical decisions about patient care and accessible by most healthcare centers around the world. With the advent of deep learning for non-medical imaging data over the past half decade, the world has quickly turned its attention to how AI could be specifically applied to medical imaging to improve clinical decision-making and to optimize workflows. Learn the fundamental skills needed to work with 2D medical imaging data and how to use AI to derive clinically-relevant insights from data gathered via different types of 2D medical imaging such as x-ray, mammography, and digital pathology. Extract 2D images from DICOM files and apply the appropriate tools to perform exploratory data analysis on them. Build different AI models for different clinical scenarios that involve 2D images and learn how to position AI tools for regulatory approval.

Project Pneumonia Detection from Chest X-Rays

Chest X-ray exams are one of the most frequent and cost-effective types of medical imaging examinations. Deriving clinical diagnoses from chest X-rays can be challenging, however, even by skilled radiologists. When it comes to pneumonia, chest X-rays are the best available method for point-of-care diagnosis. More than 1 million adults are hospitalized with pneumonia and around 50,000 die from the disease every year in the US alone. The high prevalence of pneumonia makes it a good candidate for the development of a deep learning application for two reasons: 1) Data availability in a high enough quantity for training deep learning models for image classification 2) Opportunity for clinical aid by providing higher accuracy image reads of a difficult-to-diagnose disease and/or reduce clinical burnout by performing automated reads of very common scans. In this project, you will analyze data from the NIH Chest X-ray dataset and train a CNN to classify a given chest X-ray for the presence or absence of pneumonia. First, you'll curate training and testing sets that are appropriate for the clinical question at hand from a large collection of medical images. Then, you will create a pipeline to extract images from DICOM files that can be fed into the CNN for model training. Lastly, you'll write an FDA 501(k) validation plan that formally describes your model, the data that it was trained on, and a validation plan that meets FDA criteria in order to obtain clearance of the software being used as a medical device.

Course 1: Applying AI to 2D Medical Imaging Data. cont.

LESSON TITLE	LEARNING OUTCOMES
INTRODUCTION TO AI FOR 2D MEDICAL IMAGING	• Explain what AI for 2D medical imaging is and why it is relevant.
CLINICAL FOUNDATIONS OF 2D MEDICAL IMAGING	• Learn about different 2D medical imaging modalities and their clinical applications. • Understand how different types of machine learning algorithms can be applied to 2D medical imaging. • Learn how to statistically assess an algorithm's performance. • Understand the key stakeholders in the 2D medical imaging space.
2D MEDICAL IMAGING EXPLORATORY DATA ANALYSIS	• Learn what the DICOM standard it is and why it exists. • Use Python tools to explore images extracted from DICOM files. • Apply Python tools to explore DICOM header data. • Prepare a DICOM dataset for machine learning. • Explore a dataset in preparation for machine learning.
CLASSIFICATION MODELS OF 2D MEDICAL IMAGES	• Understand architectures of different machine learning and deep learning models, and the differences between them. • Split a dataset for training and testing an algorithm. • Learn how to define a gold standard. • Apply common image pre-processing and augmentation techniques to data. • Fine-tune an existing CNN architecture for transfer learning with 2D medical imaging applications. • Evaluate a model's performance and optimize its parameters.
TRANSLATING AI ALGORITHMS FOR CLINICAL SETTINGS WITH THE FDA	• Learn about the FDA's risk categorization for medical devices and how to define an Intended Use statement. • Identify and describe algorithmic limitations for the FDA. • Translate algorithm performance statistics into clinically meaningful information that can trusted by professionals. • Learn how to create an FDA validation plan.

Course 2: Applying AI to 3D Medical Imaging Data

3D medical imaging exams such as CT and MRI serve as critical decision-making tools in the clinician's everyday diagnostic armamentarium. These modalities provide a detailed view of the patient's anatomy and potential diseases, and are a challenging though highly promising data type for AI applications. Learn the fundamental skills needed to work with 3D medical imaging datasets and frame insights derived from the data in a clinically relevant context. Understand how these images are acquired, stored in clinical archives, and subsequently read and analyzed. Discover how clinicians use 3D medical images in practice and where AI holds most potential in their work with these images. Design and apply machine learning algorithms to solve the challenging problems in 3D medical imaging and how to integrate the algorithms into the clinical workflow.

Project: Hippocampal Volume Quantification in Alzheimer's Progression

Hippocampus is one of the major structures of the human brain with functions that are primarily connected to learning and memory. The volume of the hippocampus may change over time, with age, or as a result of disease. In order to measure hippocampal volume, a 3D imaging technique with good soft tissue contrast is required. MRI provides such imaging characteristics, but manual volume measurement still requires careful and time consuming delineation of the hippocampal boundary. In this project, you will go through the steps that will have you create an algorithm that will help clinicians assess hippocampal volume in an automated way and integrate this algorithm into a clinician's working environment. First, you'll prepare a hippocampal image dataset to train the U-net based segmentation model, and capture performance on the test data. Then, you will connect the machine learning execution code into a clinical network, create code that will generate reports based on the algorithm output, and inspect results in a medical image viewer. Lastly, you'll write up a validation plan that would help collect clinical evidence of the algorithm performance, similar to that required by regulatory authorities.

LESSON TITLE	LEARNING OUTCOMES
INTRODUCTION TO AI FOR 3D MEDICAL IMAGING	• Explain what AI for 2D medical imaging is and why it is relevant.

Course 2: Applying AI to 3D Medical Imaging Data, cont.

LESSON TITLE	LEARNING OUTCOMES
3D MEDICAL IMAGING — CLINICAL FUNDAMENTALS	• Identify medical imaging modalities that generate 3D images. • List clinical specialties who use 3D images to influence clinical decision making. • Describe use cases for 3D medical images. • Explain the principles of clinical decision making. • Articulate the basic principles of CT and MR scanner operation. • Perform some of the common 3D medical image analysis tasks such as windowing, MPR and 3D reconstruction.
3D MEDICAL IMAGING EXPLORATORY DATA ANALYSIS	• Describe and use DICOM and NIFTI representations of 3D medical imaging data. • Explain specifics of spatial and dimensional encoding of 3D medical images. • Use Python-based software packages to load and inspect 3D medical imaging volumes. • Use Python-based software packages to explore datasets of 3D. medical images and prepare it for machine learning pipelines • Visualize 3D medical images using open software packages.
3D MEDICAL IMAGING — DEEP LEARNING METHODS	• Distinguish between classification and segmentation problems as they apply to 3D imaging. • Apply 2D, 2.5D and 3D convolutions to a medical imaging volume. • Apply U-net algorithm to train an automatic segmentation model of a real-world CT dataset using PyTorch. • Interpret results of training, measure efficiency using Dice and Jaccard performance metrics.
DEPLOYING AI ALGORITHMS IN THE REAL WORLD	• Identify the components of a clinical medical imaging network and integration points as well as DICOM protocol for medical image exchange. • Define the requirements for integration of AI algorithms. • Use tools for modeling of clinical environments so that it is possible to emulate and troubleshoot real-world AI deployments. • Describe regulatory requirements such as FDA medical device framework and HIPAA required for operating AI for clinical care. • Provide input into regulatory process, as a data scientist.

Course 3: Applying AI to EHR Data

With the transition to electronic health records (EHR) over the last decade, the amount of EHR data has increased exponentially, providing an incredible opportunity to unlock this data with AI to benefit the healthcare system. Learn the fundamental skills of working with EHR data in order to build and evaluate compliant, interpretable machine learning models that account for bias and uncertainty using cutting-edge libraries and tools including TensorFlow Probability, Aequitas, and Shapley. Understand the implications of key data privacy and security standards in healthcare. Apply industry code sets (ICD10-CM, CPT, HCPCS, NDC), transform datasets at different EHR data levels, and use TensorFlow to engineer features.

Project — Patient Selection for Diabetes Drug Testing

EHR data is becoming a key source of real-world evidence (RWE) for the pharmaceutical industry and regulators to make decisions on clinical trials. In this project, you will act as a data scientist for an exciting unicorn healthcare startup that has created a groundbreaking diabetes drug that is ready for clinical trial testing. Your task will be to build a regression model to predict the estimated hospitalization time for a patient in order to help select/filter patients for your study. First, you will perform exploratory data analysis in order to identify the dataset level and perform feature selection. Next, you will build necessary categorical and numerical feature transformations with TensorFlow. Lastly, you will build a model and apply various analysis frameworks, including TensorFlow Probability and Aequitas, to evaluate model bias and uncertainty.

LESSON TITLE	LEARNING OUTCOMES
EHR DATA SECURITY AND ANALYSIS	• Understand U.S. healthcare data security and privacy best practices (e.g. HIPAA, HITECH) and how they affect utilizing protected health information (PHI) data and building models. • Analyze EHR datasets to check for common issues (data leakage, statistical properties, missing values, high cardinality) by performing exploratory data analysis.
EHR CODE SETS	• Understand the usage and structure of key industry code sets (ICD, CPT, NDC). • Group and categorize data within EHR datasets using code sets.

Course 3: Applying AI to EHR Data, cont.

LESSON TITLE	LEARNING OUTCOMES
EHR TRANSFORMATIONS & FEATURE ENGINEERING	• Use the TensorFlow Dataset API to scalably extract, transform, and load datasets. • Build datasets aggregated at the line, encounter, and longitudinal(patient) data levels. • Create derived features (bucketing, cross-features, embeddings) utilizing TensorFlow feature columns on both continuous and categorical input features.
BUILDING, EVALUATING, AND INTERPRETING MODELS	• Analyze and determine biases for a model for key demographic groups by evaluating performance metrics across groups by using the Aequitas framework. • Train a model that provides an uncertainty range with the TensorFlow Probability library. • Use Shapley values to select features for a model and identify the marginal contribution for each selected feature.

Course 4: Applying AI to Wearable Device Data

Wearable devices are an emerging source of physical health data. With continuous, unobtrusive monitoring they hold the promise to add richness to a patient's health information in remarkable ways. Understand the functional mechanisms of three sensors (IMU, PPG, and ECG) that are common to most wearable devices and the foundational signal processing knowledge critical for success in this domain. Attribute physiology and environmental context's effect on the sensor signal. Build algorithms that process the data collected by multiple sensor streams from wearable devices to surface insights about the wearer's health.

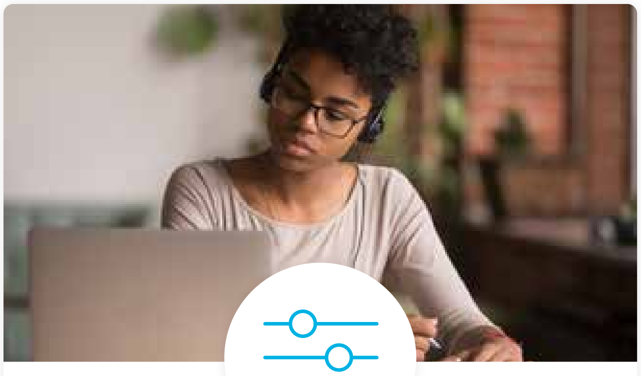
Project: Motion Compensated Pulse Rate Estimation

Wearable devices have multiple sensors all collecting information about the same person at the same time. Combining these data streams allows us to accomplish many tasks that would be impossible from a single sensor. In this project, you will build an algorithm which combines information from two of the sensors that are covered in this course -- the IMU and PPG sensors -- that can estimate the wearer's pulse rate in the presence of motion. First, you'll create and evaluate an activity classification algorithm by building signal processing features and a random forest model. Then, you will build a pulse rate algorithm that uses the activity classifier and frequency domain techniques, and also produces an associated confidence metric that estimates the accuracy of the pulse rate estimate. Lastly, you will evaluate algorithm performance and iterate on design until the desired accuracy is achieved.

Course 4: Applying AI to Wearable Device Data, cont.

LESSON TITLE	LEARNING OUTCOMES
INTRO TO DIGITAL SAMPLING & SIGNAL PROCESSING	• Describe how to digitally sample analog signals. • Apply signal processing techniques (eg. filtering, resampling, interpolation) to time series signals. • Apply frequency domain techniques (eg. FFT, STFT, spectrogram) to time series signals. • Use matplotlib's plotting functionality to visualize signals.
INTRODUCTION TO SENSORS	• Describe how sensors convert a physical phenomenon into an electrical one. • Understand the signal and noise characteristics of the IMU and PPG signals.
ACTIVITY CLASSIFICATION	• Perform exploratory data analysis to understand class imbalance and subject imbalance. • Gain an intuitive understanding signal characteristics and potential feature performance. • Write code to implement features from literature. • Recognize the danger overfitting of technique (esp. on small datasets), not simply of model parameters or hyperparameters.
ECG SIGNAL PROCESSING	• Understand the electrophysiology of the heart at a basic level. • Understand the signal and noise characteristics of the ECG. • Understand how atrial fibrillation manifests in the ECG. • Build a QRS complex detection algorithm. • Build an arrhythmia detection algorithm from a wearable ECG signal. • Understand how models can be cascaded together to achieve higher-order functionality.

Our Nanodegree Programs Include:



Pre-Assessments

Our in-depth workforce assessments identify your team's current level of knowledge in key areas. Results are used to generate custom learning paths designed to equip your workforce with the most applicable skill sets.

Dashboard & Progress Reports

Our interactive dashboard (enterprise management console) allows administrators to manage employee onboarding, track course progress, perform bulk enrollments and more.

Industry Validation & Reviews

Learners' progress and subject knowledge is tested and validated by industry experts and leaders from our advisory board. These in-depth reviews ensure your teams have achieved competency.

Real World Hands-on Projects

Through a series of rigorous, real-world projects, your employees learn and apply new techniques, analyze results, and produce actionable insights. Project portfolios demonstrate learners' growing proficiency and subject mastery.

Our Review Process

Real-life Reviewers for Real-life Projects

Real-world projects are at the core of our Nanodegree programs because hands-on learning is the best way to master a new skill. Receiving relevant feedback from an industry expert is a critical part of that learning process, and infinitely more useful than that from peers or automated grading systems. Udacity has a network of over 900 experienced project reviewers who provide personalized and timely feedback to help all learners succeed.

Vaibhav
UDACITY LEARNER

"I never felt overwhelmed while pursuing the Nanodegree program due to the valuable support of the reviewers, and now I am more confident in converting my ideas to reality."

now at

CODING VISIONS INFOTECH

All learners benefit from:

- Line-by-line feedback for coding projects
- Industry tips and best practices
- Advice on additional resources to research
- Unlimited submissions and feedback loops

How it Works

Real-world projects are integrated within the classroom experience, making for a seamless review process flow.

- Go through the lessons and work on the projects that follow
- Get help from your technical mentor, if needed
- Submit your project work
- Receive personalized feedback from the reviewer
- If the submission is not satisfactory, resubmit your project
- Continue submitting and receiving feedback from the reviewer until you successfully complete your project

About our Project Reviewers

Our expert project reviewers are evaluated against the highest standards and graded based on learners' progress. Here's how they measure up to ensure your success.

900+
Expert Project Reviewers
Are hand-picked to provide detailed feedback on your project submissions.

1.8M
Projects Reviewed
Our reviewers have extensive experience in guiding learners through their course projects.

3
Hours Average Turnaround
You can resubmit your project on the same day for additional feedback.

4.85 /5
Average Reviewer Rating
Our learners love the quality of the feedback they receive from our experienced reviewers.

UDACITY
FOR ENTERPRISE

Udacity © 2020

2440 W El Camino Real, #101
Mountain View, CA 94040, USA - HQ

For more information visit: **www.udacity.com/enterprise**