Chest x-ray interpretation is a fundamental aspect of a radiologist's training, however radiologists can undergo up to 13 years of training before being fully certified in a particular specialization. In fact, Merritt Hawkins, the largest physician search and consulting firm in the United States, reports that radiology was the 10th most recruited specialty for physician searches in 2016, up from 19th the year before. Further, the American College of Radiology tracks radiologist hiring trends, and saw a 14% increase in radiologist hires from 2016 to 2017. This drastic increase in demand for trained radiologists leaves the 2025 predicted shortage of Radiologists in the U.S. in the tens of thousands, and has resulted in a commensurate increase in telehealth radiology due to the rising costs of employment, even after accounting for COVID-19 shutdowns. The decentralization of patient care can lead to misdiagnosis of illnesses as well as delays in patient care.

To respond to this growing need, we perform an exploratory analysis using machine-learning algorithms to identify the presence of pneumonia in chest x-rays. The models we employ are easily deployable for commercial use, can be implemented without specialized computational hardware and are designed to improve healthcare KPI's such as the average treatment charge and patient wait time while reducing mistakes in treatment by accurately differentiating nominal chest x-rays from those that exhibit pneumonia. We identify three models to compare via our analysis that are well-suited to image classification problems. The first is *k*-nearest neighbors (*k*-NN), which we chose as a baseline model because x-rays exhibiting pneumonia tend to show brighter pixels in the lungs, which should allow an algorithm like *k*-NN to roughly separate the two in the high-dimensional feature space. The second model is a homemade convolutional neural network (CNN), chosen because of the great success CNN's have shown in pattern recognition, as well as their resilience to translations in the image. This CNN is composed of three convolutional layers interspersed with max pooling layers to reduce model complexity followed by three dense layers interspersed with dropout layers for regularization. Finally, the third model we use is another CNN pretrained on the CIFAR image dataset known as RESNET-152v2, which we will call the Transfer Learning (TL) model. RESNET-152v2 is a deep residual network, so-called because it consists of many stacked "residual units", which are groupings of layers that are connected to one another, along with skip connections between residual units. This architecture allows for a much larger number of layers without running aground of the vanishing gradient problem. We make use of transfer learning in our TL model by freezing the weights of RESNET-152v2 and adding several trainable layers after its outputs.

We evaluate our models using the F1 score because it is conservatively representative of the combined precision and recall performance. After training our models and scoring them on the validation set, the *k*-NN model achieves an F1 score of 0.944, the CNN achieves a score of 0.964, and the TL model achieves a score of 0.915. As a result, we choose the custom-trained CNN as our final model. This model achieves a test-sample F1 score of 0.964. We recommend our CNN model to healthcare providers because the low false negative rate means that this model will correctly identify cases of pneumonia so that patients receive the care they need. Further, the low false positive rate means that the time of staff radiologists will be used efficiently because they will not have to spend time examining x-rays of healthy patients. This model has potential for generalized high-impact applications in biomedical imaging.