In the MRNet paper, they proposed an automated system to detect anterior cruciate ligament(ACL) tears, meniscal tears and abnormalities in magnetic resonance imaging (MRI) exams of the knee. The proposed system is a deep learning model trained on the MRNet dataset. The data set is Digital Imaging and Communication in Medicine (DICOM) and clinical reports for knee MRI exams performed at Stanford Medical Center. Three types of series for each exam were used in the proposed experiment: sagittal plane T2-weighted series, coronal plane T1-weighted series, and axial plane PD-weighted series. Each series has slices ranging between 17 and 61. The dataset was split into training, tuning and validation sets, with each set containing at least 50 exams having a true label for each of: abnormality, ACL tears and meniscal tears.

The system has three models one for predicting each of ACL tears, meniscal tears and abnormalities. Each individual model is composed of three blocks of MRNet, which is a convolutional neural network that starts with a feature extractor based on AlexNet followed by an average pooling layer then a maximum pooling layer. The final layer is a dense layer with a sigmoid activation function and one output perceptron to give the output probability. Each of the three MRNet blocks predicts a probability given an input series of one type (sagittal, coronal or axial), then the three probabilities are passed through a logistic regression model to give one output probability for each examination. In the proposed paper, they applied a transfer learning approach due to the limitations in the training data size, where the weights of AlexNet are initially set to the optimal weights fitting the ImageNet database. Weights were then fine-tuned to fit the MRNet dataset, where they trained each model using binary cross-entropy loss and backpropagation algorithm.

The described experiment starts by extracting images from DICOM files then performing image pre-processing. Images were scaled to 256 x 256 pixels, converted to Portable Network Graphics (PNG) format and subjected to intensity standardization. Finally, image pixels were adjusted to the standard pixels range of PNG images. For the training set, they applied data augmentation where each example was randomly rotated, shifted and flipped horizontally. Also, the actual output of training examples was extracted from the clinical reports to be fed to the model. The input of each MRNet block has dimension s x 3 x 256 x 256, where s is a variable parameter representing the number of images in the input series and 3 is for color channels. Since the input image has no color channels, each image is replicated three times to fit the input dimension. The feature extractor is fed with the input to give a tensor of dimension s x 256 x 7 x 7 having the feature map of each image in the series, which is then passed through the average pooling layer to reduce the feature maps' dimension to s x 256. Finally, a global average is applied to all series images to give a 256-dimension vector that is passed to the final dense layer and predicts the output probability for the specified type.

The model was evaluated on an internal validation set (test set) labeled by MSK radiologists and it achieved accuracies of 0.85, 0.867 and 0.725 for general abnormalities, ACL tears and meniscus tears, respectively. The area under the receiver operating characteristic curve (AUC) for the three injuries, respectively, were 0.937, 0.965 and 0.847. Furthermore, the trained model was evaluated on an external dataset with sagittal exams of ACL injuries and it achieved 0.824 AUC.
The future work is to improve the model performance and generalize deep learning models for MRI.