SIIM CMIMI24
Oct 21-22 | Boston, MA
Conference On Machine Intelligence In Medical Imaging

A Technical Exploration of Bone Age Prediction with Machine Learning Regression Algorithms: Conformal Prediction

Shahriar Faghani, MD, Mayo Clinic; Kelly Horst, MD; Bradley J. Erickson, MD, PhD, CIIP, FSIIM

Introduction/Background

Regression models usually provide a point estimate rather than an interval that reflects the uncertainty of the prediction. While some uncertainty quantification methods can be applied to regression algorithms to obtain prediction intervals, most lack statistical guarantees for these intervals. Conformal prediction, a post-hoc method, stands out by offering statistical guarantees. We aimed to apply conformal prediction to bone age prediction in the pediatric population.

Methods/Intervention

A multimodal deep learning model was developed to estimate the skeletal age based on left-sided hand radiographs. This model consists of the DenseNet121 architecture as an imaging feature extractor and two shallow neural networks: one for transforming patient sex into the imaging feature space and the other for combining all features. The 2017 RSNA bone age dataset was utilized, with 12,611 training images, 700 and 725 reserved for the calibration and validation set, respectively. There are 5,778 female and 6,833 male subjects, ranging from 1 to 228 months in age, with a mean bone age of 127.2 months and a standard deviation of 41.7 months. Resizing to 512 x 512, foreground cropping, normalization of image intensities based on training dataset statistics to zero mean and unit standard deviation, histogram shift, flipping, affine transformations, and Gaussian noise addition were used to enhance model generalizability. Quantile regression loss with the 5th and 95th percentile provided at least 90% coverage for interval predictions. A calibration set was used to calibrate the predicted percentiles for each image based on the conformal procedure. Mean absolute error (MAE), and mean prediction interval (MPI) were calculated.

Results/Outcome

An MAE of 5.6, an MPI of 3.7, and a prediction interval coverage of 97% were achieved, indicating that 97% of the observations fall within the predicted intervals.

Conclusion

These intervals allow for the individualization of the result uncertainty quantification and may be helpful to include in individual patient reports. This measurement may also assist in identifying individual outliers, which may need further follow-up.

Statement of Impact

To the best of our knowledge, this is the first application of conformal prediction in radiologic image-level regression tasks, that could be served as a template for other clinical challenges.

Keywords

Deep learning; Uncertainty quantification; Regression; Pediatric radiology; MSK radiology; Conformal prediction