Enhancing Diagnostic Brain CT Referral Justification with AI

Diagnostic imaging is crucial in modern medicine for quick and accurate diagnosis and monitoring of various conditions. However, its overuse leads to unnecessary expenses, increased radiation exposure, and potential harm. The global use of imaging like CT and MRI has risen significantly, with a notable proportion being unnecessary. The FDA estimates a substantial percentage of CT scans in the US are unjustified. Efforts to address this issue include implementing diagnostic pathways, specialist involvement, and post-imaging audits. Challenges such as financial constraints, equipment availability, and adherence to guidelines hinder these efforts. Retrospective audits can improve guideline compliance but are often limited by resources. Real-time justification processes and clinical decision support systems using AI show promise in enhancing guideline implementation. Recent studies used AI techniques like NLP and machine learning to automate justification analysis of imaging referrals. To tackle inappropriate CT scans effectively, a study recently published in European Radiology developed an AI-based interpreter to standardize justification practices across clinical sites, focusing initially on brain CT scans. This approach aimed to compare AI performance with human experts in enhancing imaging justification standards.

Multi-Site Analysis of Brain CT Referrals

Three tertiary referral hospitals, consisting of two private and one public institution, provided anonymized data from brain CT scan referrals for adult patients throughout 2020 and 2021. The data included patient demographics (gender, age) and unstructured clinical indications. Referrals were randomly shuffled and divided into groups of approximately 1000 each per hospital site, with one group randomly selected from each site. Duplicate and inadequate referrals lacking clinical indications were excluded. Two radiographers with 6 and 8 years of experience initially assessed justification, followed by review from two consultant radiologists with 11 and 15 years of experience in cases of disagreement. Justification decisions adhered to iGuide clinical imaging guidelines, facilitated by the xWave CDS platform for accessing recommendations specific to clinical indications. Final justification labels were determined by majority vote among annotators, with consensus reached in cases of ties or insufficient clinical evidence.

Advanced Data Processing and Model Training for Brain CT Referral Analysis

Data preprocessing involved tokenizing unstructured clinical indications using the Natural Language Toolkit (NLTK) word tokenizer, normalizing tokens, and spell-checking using the Enchant algorithm with a custom medical dictionary. Stop words were filtered out, and abbreviations were expanded to improve text quality without altering the clinical context. Bag-of-words (BoW), term frequency-inverse document frequency (TF-IDF), and sentence-level vectors generated from a pretrained clinical Word2vec model were used for feature extraction. The dataset was randomly undersampled to balance classes and split into stratified training and test sets (80% training, 20% test). Machine learning (ML) models (SVM, logistic regression, gradient-boosting ensemble) and deep learning (DL) models (Bi-LSTM, MLP) were trained and evaluated on the test set. The Bi-LSTM model, chosen for its superior performance in handling free-text clinical data, utilized 128-dimensional word embeddings and a single Bi-LSTM layer with 100 units. Training stopped based on validation loss criteria, with 50 epochs set for training. Evaluation metrics included accuracy and the F1 score for each label, comparing model performance with inter-rater agreement among radiographers. The study focused on enhancing the justification process for brain CT scans across multiple clinical sites, leveraging AI and clinical decision support to improve adherence to imaging guidelines and reduce unnecessary scans.

Classifiers showed minimal false predictions compared to disagreements among radiographers

Among the 11,090 initial referrals from three hospitals (site A: 2651, site B: 4351, site C: 4088), 2958 referrals underwent justification analysis after excluding 42 disregarded cases. Two radiographers initially disagreed on 946 (32.0%) referrals, which were subsequently reviewed by two consultants. Consensus was reached for 839 (88.7%) referrals, with the remaining 107 (11.3%) resolved through further discussion. Among the reviewed referrals, radiologists disagreed on 274 (29.0%). Site A, a private facility, had the lowest rate of justified scans (45.0%), followed by site B (69.8%) and public site C (79.4%). Overall, 8.1% of referrals were unjustified, with sites A and B contributing 58.8% and 30.3% respectively. Potentially justified referrals constituted 27.4%, predominantly from sites A and B (79.0%).

Common inappropriate indications for CT included dizziness/vertigo and long-lasting headaches without new features, while appropriate indications included stroke/transient ischemic attack and head injury. MRI was often identified as a suitable alternative to CT.

In machine learning analysis, the gradient-boosting classifier with Bag-of-Words embeddings achieved the highest performance, with 94.4% accuracy and a macro F1 score of 0.94. TF-IDF-based support vector machines also performed well, achieving 93.7% accuracy and an identical macro F1 score. Deep learning models like Bi-LSTM achieved 92.3% accuracy and a macro F1 of 0.92, outperforming MLP models which ranged between 90% accuracy and 0.91 macro F1. The classifiers showed minimal false predictions compared to disagreements among radiographers.

Disparities in Referral Justification for Brain CT Scans Between Machine Learning and Hospital

Overall, the study highlighted disparities in justification rates across hospitals and demonstrated effective machine learning approaches in improving the accuracy of referral justifications for brain CT scans. The study underscores the critical role of justification in both radiation protection and efficient healthcare resource utilization, particularly in the context of CT imaging which contributes significantly to population radiation dose. Analysis revealed disparities among hospitals, with private facilities (sites A and B) showing lower rates of justified scans compared to a public site (site C), suggesting potential financial motivations influencing referral practices. The research demonstrated that machine learning (ML) and deep learning (DL) models can effectively automate the analysis of CT referrals based on unstructured clinical indications, outperforming human reviewers in some instances.

Navigating Variability in Clinical Referrals with AI-driven Justification Analysis

The variability in referral quality and style among different referrers and institutions was evident, with referrals often containing slang, abbreviations, misspellings, or overly detailed clinical histories. Ambiguities in clinical presentations posed challenges during justification audits, contributing to disagreements among radiographers and consultants. ML and DL models, trained on such heterogeneous data, achieved high accuracy in classifying referrals into justified, unjustified, and potentially justified categories despite inherent dataset imbalances.

The study highlighted the need for robust datasets and sophisticated models to handle the complexity and variability of clinical language effectively. It noted that traditional feature extraction techniques like Bag-of-Words (BoW) and TF-IDF performed well, while DL models such as Bi-LSTM showed promise but required careful handling due to dataset limitations and potential overfitting. The research emphasized the potential of AI-driven natural language processing (NLP) to enhance the efficiency and accuracy of referral justification processes, thereby integrating clinical guidelines more consistently across diverse healthcare settings.

Looking forward, future work should explore transfer learning and larger, more balanced datasets to further refine AI models in interpreting clinical indications. This advancement could lead to improved collaboration between AI systems and human experts, enhancing the overall quality and appropriateness of diagnostic imaging referrals. Ethical and legal considerations must accompany such implementations to ensure patient safety and regulatory compliance as AI continues to evolve in clinical decision support.

Source Credit: European Radiology
Image Credit: iStock

Published on : Thu, 27 Jun 2024