Abstract

TITLE: Computer-Assisted Coding and Documentation in a Large Hospital in Milan, Italy

Introduction

In 2024, a hospital conducted a project aimed at enhancing the quality of clinical data coding through the development and deployment of a solution based on natural language processing (NLP). The project focused on training the technology using semantic analysis of the hospital’s clinical documentation, enabling it to provide higher-quality coding performance for medical records.

Methods

The project began with a proof of concept (POC) based on NLP technology, processing 2,387 discharge letters to extract additional diagnosis and procedure codes. It culminated in more accurate Electronic Health Records (EHRs), aligned with local coding standards. The POC took place in two phases. The initial phase involved an in-depth analysis of the hospital's clinical documentation to identify key semantic patterns and terminologies relevant to the coding process. Leveraging advanced NLP technology, the solution was iteratively trained to understand and interpret clinical data with high precision. This phase concluded when the system's performance metrics, specifically precision and sensitivity, met acceptable thresholds, ensuring the solution’s reliability and readiness for operational use. In the second phase, the trained solution was deployed across three departments. During this phase, further training continued, supported by feedback from clinicians and coders.

Results

Baseline results were measured by comparing the codes registered in the medical records to those in the clinical documentation. The most significant result from the NLP technology, which was unexpected by the hospital, was the improvement in coding quality by extracting a consistent number of missing principal diagnosis and procedure codes from the medical records. Another notable advancement was the improvement in NLP performance, facilitated by the review of semantic rules by coders and clinicians. Their feedback played a crucial role in refining the solution, enabling it to adapt to specific workflows and address nuances in the hospital’s documentation practices. The implementation process also included user training sessions to familiarize staff with the tool’s features and ensure its seamless integration into existing workflows. The software served as a decision-support system, providing real-time coding suggestions and flagging potential errors or inconsistencies in the coding of discharge letters.

Discussion/Conclusions

The efficiency and accuracy of coding can significantly benefit from incorporating natural language processing (NLP) technology into the workflow. Even in hospitals with high perceived and observed quality of coding, there are still areas for improvement. This project highlights the importance of tailoring technological innovations to meet the specific needs of healthcare providers, ensuring practical usability and measurable outcomes. Looking ahead, the scalability of the solution offers opportunities for broader adoption. In 2025, a new phase will be launched to extend the NLP technology across additional

departments, further contributing to the standardization and quality of clinical data coding. This phase will also trigger up-to-date regional coding edits, warnings, and measure the impact on professional time, case-mix index, and reimbursements.