question @471 112 views

Actions

Midterm Prep: 3. Healthcare data, Health Data standards

Howdy everyone,

In preparation for the midterm, I've put together some notes on the various topics, drawing from @416 and the practice midterm.

Going through these notes and completing the practice midterm & quiz has definitely boosted my confidence! Here, I'm happy to share these notes with you all.

Let's ace the midterm together!

Feel free to contribute and add your own insights to these notes as well.

Best regards,
Your somewhat helpful AI bot, Darin

exam

Edit good question | 1 Updated 2 days ago by Darin Zhen

S **the students' answer,** *where students collectively construct a single answer*

Actions

Healthcare data

Healthcare data for deep learning encompasses a wide range of information collected from various sources within the healthcare system. This data is used to train deep learning models to assist in medical diagnosis, prognosis, treatment planning, and patient management.

Here are some key aspects of healthcare data for deep learning:

1. **Types of Healthcare Data:**

- **Electronic Health Records (EHR):** EHRs contain structured and unstructured data such as patient demographics, medical history, diagnoses, medications, laboratory results, and imaging reports.
- **Medical Images:** This includes various types of medical imaging data such as X-rays, MRI scans, CT scans, ultrasound images, and histopathology slides.
- **Genomic Data:** Genomic data includes DNA sequences, gene expression profiles, and genetic variations associated with diseases and drug responses.
- **Medical Sensor Data:** Data from wearable devices, medical sensors, and monitoring devices provide continuous physiological measurements such as heart rate, blood pressure, temperature, and activity levels.
- **Clinical Notes and Reports:** Free-text clinical notes, discharge summaries, and radiology reports contain valuable information that can be used for natural language processing tasks.

2. **Challenges:**

- **Data Privacy and Security:** Healthcare data is sensitive and subject to privacy regulations such as the Health Insurance Portability and Accountability Act (HIPA the United States. Data anonymization and encryption techniques are crucial to protect patient privacy.
- **Data Heterogeneity:** Healthcare data is often heterogeneous, coming from different sources, formats, and standards. Integrating and preprocessing such divers types is challenging.
- **Data Imbalance:** Some medical conditions are rare, leading to imbalanced datasets where one class significantly outweighs the others. Addressing class imbal essential to prevent biased model predictions.
- **Data Quality:** Ensuring the quality and accuracy of healthcare data is critical for training reliable deep learning models. Data may contain errors, missing values, inconsistencies that need to be addressed during preprocessing.

3. **Applications of Deep Learning in Healthcare:**

- **Medical Imaging:** Deep learning models can analyze medical images to detect abnormalities, segment organs or lesions, and assist radiologists in diagnosis. F example, deep learning has been used for the early detection of diseases like cancer and Alzheimer's disease from imaging data.
- **Clinical Decision Support:** Deep learning models can aid clinicians in making more accurate diagnoses, predicting disease progression, and recommending personalized treatment plans based on individual patient data.
- **Drug Discovery and Development:** Deep learning techniques are used in drug discovery for identifying potential drug candidates, predicting drug-target intera and optimizing drug design.
- **Healthcare Management:** Deep learning models can analyze EHR data to improve hospital operations, predict patient outcomes, and identify at-risk populatio preventive interventions.

4. **Ethical Considerations:**

- **Bias and Fairness:** Deep learning models trained on healthcare data may exhibit biases, leading to disparities in healthcare delivery. Ensuring fairness and equi model predictions is essential to avoid exacerbating existing disparities.
- **Transparency and Interpretability:** Deep learning models are often considered "black boxes," making it challenging to understand how they arrive at their pred Ensuring model transparency and interpretability is crucial for gaining trust from clinicians and patients.

Healthcare data presents unique challenges and opportunities for deep learning applications, and careful consideration of these factors is essential for the successful development and deployment of deep learning models in healthcare settings.

Healthcare data standards

Healthcare data standards play a crucial role in ensuring interoperability, consistency, and quality of data used in deep learning applications within the healthcare dom These standards help facilitate data exchange, integration, and analysis across different systems and organizations.

Some of the key healthcare data standards relevant to deep learning applications include:

1. International Classification of Diseases (ICD) and Current Procedural Terminology (CPT):

ICD and CPT are standardized code sets used for classifying diseases, medical diagnoses, and procedures. They provide a common language for encoding and categ clinical information, which facilitates data exchange, billing, and reporting across healthcare systems.

2. Logical Observation Identifiers Names and Codes (LOINC):

LOINC is a standard for identifying and exchanging laboratory and clinical observations, including test results, measurements, and assessments. It provides a universa system for naming and coding observations, ensuring consistency and interoperability across different healthcare settings.

3. SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms):

SNOMED CT is a comprehensive clinical terminology system that provides standardized codes and concepts for describing clinical findings, procedures, and other healthcare-related concepts. It enables semantic interoperability by supporting precise and unambiguous representation of clinical data.

Edit | undo thanks | 2 | Updated 2 days ago by Darin Zhen

followup discussions *for lingering questions and comments*

Resolved | Unresolved | @471_f1

Darin Zhen 2 days ago

Notes – Old Book: Health Data:

- Electronic health records (EHRs) have grown rapidly, with nearly all US hospitals adopting certified EHR systems. EHRs contain diverse health data including structured data like diagnoses, procedures, medications, lab tests, as well as unstructured data like clinical notes.
- Structured EHR data uses standard codes like ICD for diagnoses, CPT for procedures, and LOINC for labs. These codes are hierarchical and sparse.
- Unstructured EHR data includes clinical notes documenting patient encounters. Notes have different types like progress notes, discharge summaries, radiology reports. They are difficult to analyze due to high dimensionality and privacy concerns.
- Other health data include continuous signals like ECG, EEG; medical images like x-rays, CT, MRI; and biomedical data for drug discovery containing molecular structures and knowledge graphs.
- Key challenges in analyzing health data include sparsity, lack of labels, noise, privacy concerns, and effective use of domain knowledge encoded in standards like SNOMED-CT and UMLS.

helpful! | 1

Reply to this followup discussion

Resolved | Unresolved | @471_f2

Darin Zhen 2 days ago

Lec02a health data

- Electronic health records (EHRs)
- Medical claims
- Clinical notes
- Medical literature
- Continuous signals like ECG
- Imaging data like X-rays, CT scans
- Medical ontologies
- Clinical trial data
- Drug discovery data

helpful! | 1

Reply to this followup discussion

Resolved Unresolved @471_f3

Darin Zhen 2 days ago

Lec02b health data standard

There are several major health data coding and terminology standards, including:

- ICD (International Classification of Diseases) - used to classify diseases and health conditions, maintained by WHO. ICD-10 is the latest version with over 140,000 codes.
- CPT (Current Procedural Terminology) - used to code medical procedures, maintained by the American Medical Association. Has category I codes for widely performed procedures.
- LOINC (Logical Observation Identifiers Names and Codes) - standard for laboratory and clinical observations, maintained by the Regenstrief Institute.
- NDC (National Drug Code) - registered by FDA, used throughout drug supply chain.
- SNOMED CT (Systematized Nomenclature of Medicine) - comprehensive clinical terminology maintained by IHTSDO. Encodes meanings and relationships between health concepts.
- UMLS (Unified Medical Language System) - integrates existing health terminology standards into a comprehensive thesaurus. Maintained by US National Library of Medicine.

These standards aim to enable exchange of health data, support analytics, and improve interoperability between health systems. Proper use of standards is crucial for recording, analyzing and sharing health data effectively.

helpful! | 1

Reply to this followup discussion

Start a new followup discussion

Compose a new followup discussion