







Article

Monitoring the Epidemiology of Otitis Using Free-Text Pediatric Medical Notes: A Deep Learning Approach

Corrado Lanera ¹, Giulia Lorenzoni ¹, Elisa Barbieri ², Gianluca Piras ¹, Arjun Magge ³, Davy Weissenbacher ³, Daniele Donà ², Luigi Cantarutti ⁴, Graciela Gonzalez-Hernandez ³, Carlo Giaquinto ^{2,4} and Dario Gregori ^{1,*}

- ¹ Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, 35131 Padova, Italy; corrado.lanera@unipd.it (C.L.); giulia.lorenzoni@unipd.it (G.L.)
- ² Division of Pediatric Infectious Diseases, Department for Woman and Child Health, University of Padova, 35128 Padova, Italy; elisa.barbieri@unipd.it (E.B.); daniele.dona@unipd.it (D.D.); carlo.giaquinto@unipd.it (C.G.)
- ³ Health Language Processing Center, Institute for Biomedical Informatics at the Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; amaggera@asu.edu (A.M.); dweissen@pennmedicine.upenn.edu (D.W.); gragon@pennmedicine.upenn.edu (G.G.-H.)
- ⁴ Società Servizi Telematici—Pedianet, 35100 Padova, Italy; l.cantarutti@sosepe.com
- * Correspondence: dario.gregori@unipd.it; Tel.: +39-049-8275384; Fax: +39-02-700445089

Abstract: Free-text information represents a valuable resource for epidemiological surveillance. Its unstructured nature, however, presents significant challenges in the extraction of meaningful information. This study presents a deep learning model for classifying otitis using pediatric medical records. We analyzed the Pedianet database, which includes data from January 2004 to August 2017. The model categorizes narratives from clinical record diagnoses into six types: no otitis, non-media otitis, non-acute otitis media (OM), acute OM (AOM), AOM with perforation, and recurrent AOM. Utilizing deep learning architectures, including an ensemble model, this study addressed the challenges associated with the manual classification of extensive narrative data. The performance of the model was evaluated according to a gold standard classification made by three expert clinicians. The ensemble model achieved values of 97.03, 93.97, 96.59, and 95.48 for balanced precision, balanced recall, accuracy, and balanced F1 measure, respectively. These results underscore the efficacy of using automated systems for medical diagnoses, especially in pediatric care. Our findings demonstrate the potential of deep learning in interpreting complex medical records, enhancing epidemiological surveillance and research. This approach offers significant improvements in handling large-scale medical data, ensuring accuracy and minimizing human error. The methodology is adaptable to other medical contexts, promising a new horizon in healthcare analytics.

Keywords: otitis; machine learning; text mining; electronic medical record; real-world data



Citation: Lanera, C.; Lorenzoni, G.; Barbieri, E.; Piras, G.; Magge, A.; Weissenbacher, D.; Donà, D.; Cantarutti, L.; Gonzalez-Hernandez, G.; Giaquinto, C.; et al. Monitoring the Epidemiology of Otitis Using Free-Text Pediatric Medical Notes: A Deep Learning Approach. *J. Pers. Med.* **2024**, *14*, 28. <https://doi.org/10.3390/jpm14010028>

Academic Editor: Juergen Hahn

Received: 30 November 2023

Revised: 20 December 2023

Accepted: 21 December 2023

Published: 25 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Otitis media (OM) is one of the most prevalent infectious diseases in childhood, causing high healthcare resource utilization, especially in primary care settings [1]. Moreover, the indiscriminate use of antibiotics for the treatment of OM has been recorded, contributing to an increase in the burden of antibiotic resistance [2]. In recent years, it has been suggested that introducing pneumococcal conjugate vaccines has resulted in a change in OM epidemiology and antibiotic prescriptions [3]. Vaccine administration seems to result in a reduction in pneumococcal Acute OM (AOM); however, the efficacy on all-cause AOM is still debated, so OM continues to represent a severe burden from a public health perspective, justifying the need for the strict epidemiological surveillance of the phenomenon [4,5]. The regular monitoring of OM enables the identification of patterns and

trends in incidence, informing prevention strategies and healthcare policies. Moreover, it aids in understanding the etiology and risk factors associated with otitis, which is essential for developing targeted interventions. Such surveillance also facilitates the early detection of emerging resistant pathogens, guiding appropriate antibiotic use. Ultimately, the effective surveillance of pediatric otitis supports public health initiatives, reduces healthcare burdens, and contributes to the overall well-being of children.

Unfortunately, studying the epidemiology of OM can be challenging because this type of disease is treated mainly in an outpatient setting, resulting in difficulties in accessing data collected by pediatricians. In Italy, the main challenge is represented by using the International Classification of Disease (ICD) codes for classifying patients' diagnoses, which is not mandatory for pediatricians working in an outpatient setting. Not least, using narratives to report data collected (e.g., for reporting clinical notes) is still common [6].

In this context, narratives can be considered a valuable source of information. In a recent study, integrating free-text clinical notes with structured information enhanced the diagnostic accuracy of acute respiratory infections encountered during outpatient visits [7]. It is worth noting that exploiting narratives can be demanding, both in terms of labor resources and costs. An automatic machine learning approach to the problem is appealing, given machine learning's potential to effectively use all the available textual information to detect the information reported in narratives. Furthermore, a deep learning approach can effectively take advantage of data that are orders of magnitude larger than classical shallow machine learning models [8].

The application of automatic free-text classification in the field of pediatric infection epidemiology presents substantial opportunities for advancement. Utilizing deep learning for the diagnosis and monitoring of otitis in children has the potential to revolutionize data collection and analysis in epidemiological studies. Firstly, deep learning enables the efficient processing of large volumes of unstructured data, such as clinical notes and medical reports, which is particularly advantageous in areas with extensive and complex datasets [9]. Secondly, automatic classification enhances accuracy and consistency in case categorization. Traditionally, this task requires significant human effort, making it prone to errors and variability. Deploying deep learning models reduces the risk of misclassification and ensures greater uniformity in data interpretation. Thirdly, automation allows for the real-time monitoring of epidemiological trends. This is crucial for the prompt identification of outbreaks or shifts in infection patterns, enabling timely and targeted interventions [10]. Additionally, the ability to analyze free text paves the way for utilizing unconventional data sources like social media and online platforms, which can provide supplementary insights into infection spread dynamics. Finally, deep learning models can aid in personalizing care. By analyzing detailed clinical narratives, it is possible to identify patterns specific to patient groups, contributing to more targeted and effective medicine for pediatric infections. This approach is particularly valuable in contexts where individual variations significantly influence infection outcomes, as in the case of otitis in children.

The present work proposes an automatic deep learning system trained using data from Pedianet to classify otitis from outpatient clinical records into six mutually exclusive categories: no otitis, otitis (not media or acute), OM (not acute), AOM, AOM with perforation, or recurrent AOM (when explicitly stated by the pediatrician in the electronic health record).

2. Materials and Methods

The data used for the present work came from the Pedianet database, which contains information on 6,903,035 visits of 216,976 children (collected by 144 family pediatricians from 1 January 2004 to 23 August 2017). Pedianet [11] is an Italian pediatric general practice research database. It contains information on the reason for the visit, health status (according to the Guidelines of Health Supervision of the American Academy of Pediatrics), personal details, growth parameters, diagnosis, and clinical details (free text or the ICD-9), prescriptions (pharmaceutical prescriptions identified by the Anatomical-Therapeutical

Chemical code, specialist appointments, diagnostic procedures, hospital admissions), and outcome data for the children. Pediatricians can access the system using standard software (JuniorBit®, Version 7.2.7) during routine patient care, and then, data are anonymized and sent monthly to a centralized database in Padova for validation. The data used in the present study for the classification task included textual information.

Records relevant to the classification were identified from the main database through a search string similar to the one used by Barbieri et al. [2] but looking at all the free-text fields. The string was built to include various potential typographical errors and abbreviations (Table S1, Supplementary Materials).

We split the records in the main database into three main sets: the training, validation, and test sets. Each dataset was obtained through a sampling strategy that ensured the same proportion of patients per pediatrician was maintained as in the main database; at least 500 records were included in each dataset, and at least one record for each pediatrician was included. The training set was derived from historical records (2004–2007), while the validation and test sets were derived from records reported in the main database after 2007.

2.1. Gold Standard

The classification of the otitis cases (gold standard) was performed by two independent evaluators (experts in pediatric care). The otitis cases were classified according to six mutually exclusive classes: no otitis, otitis (not acute), OM (not acute), AOM (without tympanic membrane perforation or recurrence), AOM with tympanic membrane perforation, and—recurrent AOM. For recurrent AOM cases, the definition used was the one of Goycoolea: “the condition in a child is defined as having at least three episodes of AOM in a period of six months, or four or more episodes in 12 months”, or with an explicit statement of the pediatricians which mark the case as recurrent [12].

Disagreements between the two independent evaluators were solved by a third independent reviewer who was specialized in infectious diseases. The agreement between the reviewers was then evaluated using the weighted Cohen’s kappa.

2.2. Data Pre-Processing

To process the free text in our study, we employed the fastText algorithm [13] with a skip-gram architecture on our main database, including various medical texts, to create dense word vectors in a 300-dimensional space. This resulted in a dictionary linking each word to a 300-dimensional vector, totaling 122,591 entries.

For data preparation, we merged all text fields from each medical visit into a single text stream using “SEP” as a separator between fields and replacing all numbers with “NUM” to simplify the dataset. Our embedding dictionary included an “OOV” token for future out-of-vocabulary words, represented by a small random vector.

Each network input was a 2-dimensional tensor, with dimensions representing batch size and a fixed word count per record (1000 words), using padding or truncation as needed. Based on our embedding dictionary, the first network layer transformed this input into the proper tensor.

2.3. Model Development

In the initial training stage, we utilized the training dataset to identify the optimal set of parameters for each model architecture. Subsequently, the most effective model from each architecture was retrained, incorporating both the original training set and a randomly selected subset of the validation set comprising 300 records. Following this, the validation set was employed to fine-tune these trained networks. This involved assessing their performance with various hyperparameter configurations to determine the most effective combination. The ultimate selection of the best models, along with their combined ensemble, was then evaluated on the test set. The performance was evaluated by computing the accuracy, balanced precision, balanced recall, and balanced F1 measure.

2.4. Architectures Employed for Model Development

Several deep learning architectures were explored and tuned on the validation set, and an ensemble model was constructed based on them [14]. The Adam optimizer trained all the networks to minimize the average training cross-entropy loss function among the batches [15].

We explored different architectures. Common parts of all of them were provided as inputs, the first hidden layer (i.e., the embedding) and the output layer. The output layer had six neurons to represent all the possible classes. It was activated by the logit function and processed by the softmax function.

We applied batch normalization after each hidden layer to maintain the control of both exploding and vanishing gradient events [16]. To avoid overfitting, we considered a dropout layer, i.e., a layer that randomly ignored a random set of neurons given a rate after each hidden layer once batch-normalized. For the embedding layer, we considered a dropout ratio of 0.2, while for the others, we explored two ratios: 0.5 and 0.7 [17]. With regard to the batch size, for each network, we explored two options: $M = 8$ or $M = 16$ [18].

The architecture explored was the following benchmark architecture (i.e., a simple embedding: “0”) plus four others of increasing complexity:

0. Simple embedding: The only hidden layer was the embedding layer.
1. Single kernel convolutional neural network (CNN): After the embedding layer, we attached a single convolutional layer.
2. Sequential single kernel CNN: After the embedding layer, we attached a sequence of two convolutional layers.
3. Multiple parallel kernel CNN: After the embedding layer, we attached a single concatenation of multiple convolutional layers.
4. Deep multiple parallel kernel CNN: After the embedding layer, we attached a sequence of two distinct concatenations of multiple convolutional layers.

The final ensemble model comprised the four networks described (simple embedding excluded), considering the mean of all their probability predictions for each class estimated by their output layer before applying their softmax activation function. The same softmax was subsequently applied to determine the class assigned to the record by the final ensemble.

We ran all the computations on an Ubuntu 18.04.3 LTS GNU/Linux 4.15.0-58-generic x86_64 virtualized server in the Unit of Biostatistics, Epidemiology, and Public Health at the University of Padova, equipped with 16 cores from Intel® Xeon® CPUs E5-2640 v4 @ 2.40 GHz, and 96 GiB-RAM. We implemented all the networks and codes for the analyses in R (v3.6.1), powered by the Keras (v2.2.4.1.9001) R interface to the TensorFlow (v1.14) backend, built from source code enabling the usage of the Intel® AVX set of instruction extensions. To learn word representation, fastText (v0.9.1) was used. Diagrams for the networks were automatically drawn from the Keras models trained using Netron (v3.3.5). All the developments and code were tracked on a GitHub repository that is available to the public at www.github.com/UBESP-DCTV/limpido (accessed on 22 December 2023).

3. Results

Figure 1 presents this study’s workflow.

The original Padianet dataset included 6,903,035 records corresponding to 216,976 children. After identifying the relevant cases, the main dataset comprised 297,373 records corresponding to 99,896 children and 142 pediatricians. The children in the main database were slightly more likely to be male (52.2%).

The training set (Table 1) included 4926 records corresponding to 4475 children. The validation set included 723 visits corresponding to 718 children, and the test set included 880 visits corresponding to 873 children. The proportion of children of the male gender was constant among the three datasets and in line with that in the main dataset (~52%).

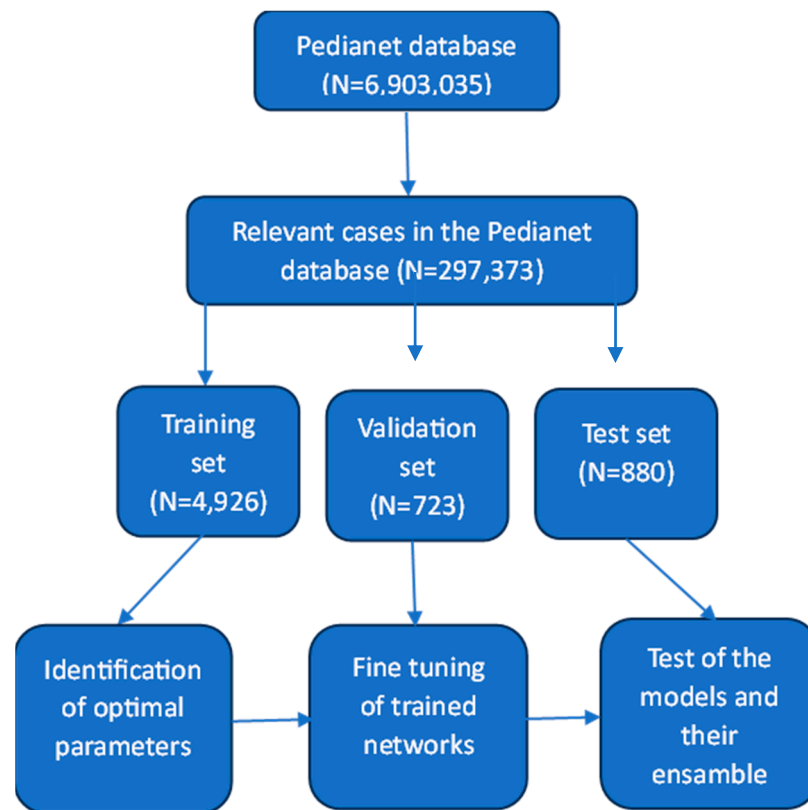


Figure 1. Study workflow.

Table 1. Characteristics (number of visits, number of children, number of pediatricians, gender of children) of the training, validation, and test sets used for model development and testing. For categorical variables, data are absolute numbers (percentages).

	Training Set	Validation Set	Test Set
Visits	4926	723	880
Pediatricians	138	142	142
Children	4475	718	873
Gender: Male	2349 (52.5%)	377 (52.5%)	463 (53.0%)
Females	2078 (46.4%)	341 (47.5%)	410 (47.0%)

Figure 2 presents the distribution of diagnostic classes according to the gold standard classification in the three datasets. Cases of AOM without perforation represented the highest proportion in all three datasets, ranging from 44% to 48%. The agreement between the independent reviewers was very good, with a weighted Cohen’s Kappa value of 0.89.

Model Performance

Tables S2–S5 (Supplementary Materials) report the performances on the validation set for the models 1–4. The validation performance was excellent, with accuracy values of at least 0.9. Table 2 presents the confusion matrix for the final ensemble model in the test set. Thirty were the cases misclassified by the ensemble model. One of the most common errors included no otitis being incorrectly predicted as otitis. A detailed review by a pediatrician of the 30 records incorrectly classified by the ensemble model revealed that these misclassifications often occurred due to negations related to otitis (four instances) or references to other doctors’ diagnoses, predispositions, or uncertain cases. Similarly, perforation was frequently misclassified as AOM, which was observed in six cases where doctors described eardrum perforation using atypical terminology.

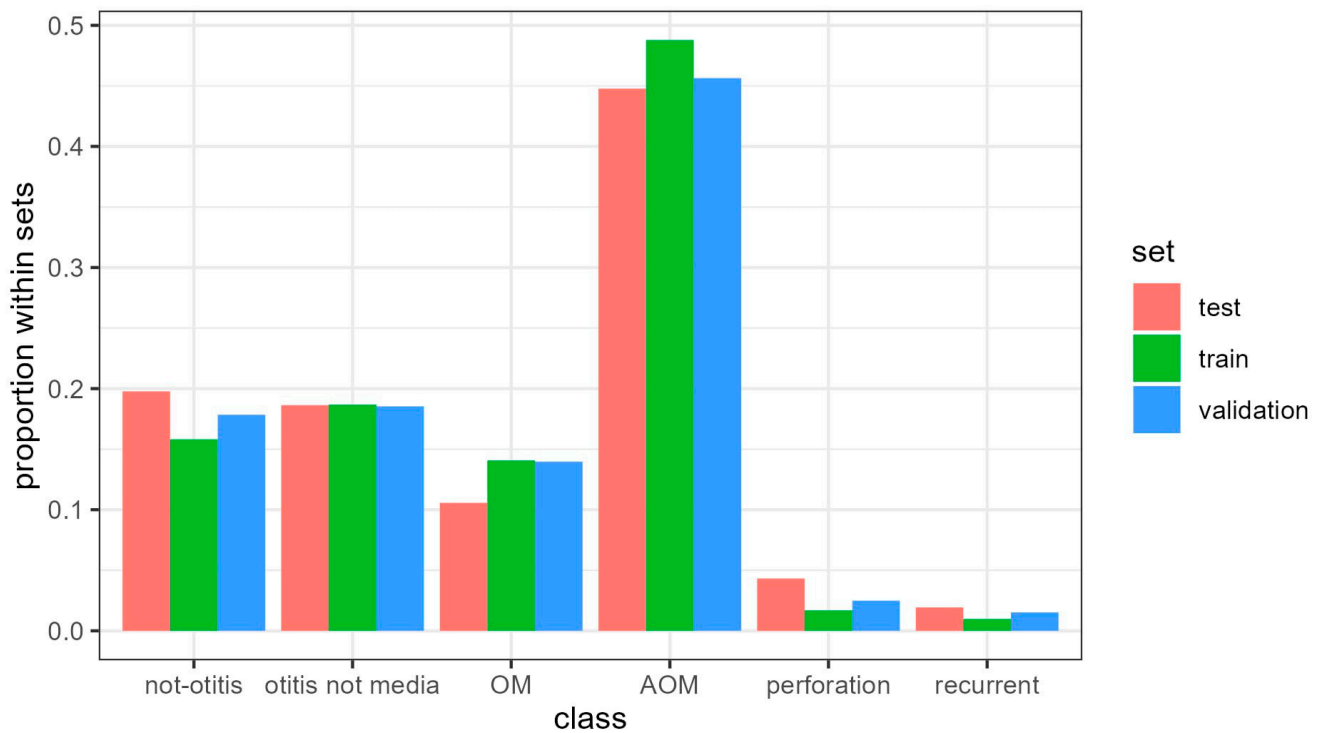


Figure 2. Proportion of diagnostic classes according to the gold standard classification in the three datasets used for model development.

Table 2. Confusion matrix for the classes predicted by the ensemble model (row) and according to the gold standard (columns).

Predicted \ Gold	No Otitis	Otitis Not Media	OM	AOM	Perforation	Recurrent	Sum
No otitis	155	0	2	0	0	0	157
Otitis not media	7	168	7	0	0	0	182
OM	1	0	101	1	1	0	104
AOM	2	1	1	389	6	0	399
Perforation	0	0	0	1	28	0	29
Recurrent	0	0	0	0	0	9	9
Sum	165	169	111	391	35	9	880

The performances on the test set for the best model selected and their ensembles are reported in Table 3. The performance of the single models was generally very good. The accuracy ranged from 81.70 to 96.59. Similarly, the balanced precision ranged from 84.51 to 96.95, and the balanced F1 measure ranged from 75.75 to 95.86. Taking the performance metrics together, the ensemble model outperformed the other architectures, with values of 97.03, 93.97, 96.59, and 95.48 for balanced precision, balanced recall, accuracy, and balanced F1 measure, respectively.

Table 3. Performances on the test set evaluated on the best model of each architecture and on the ensemble model.

Selected Network	Balanced Precision	Balanced Recall	Accuracy	Balanced F1
Simple embedding	84.51	68.63	81.70	75.75
Single kernel	92.60	91.87	94.66	92.23
Sequential CNN	95.94	81.26	93.64	87.99
Parallel CNN	96.95	94.78	96.59	95.86

Table 3. Cont.

Selected Network	Balanced Precision	Balanced Recall	Accuracy	Balanced F1
Deep CNN	96.38	93.36	96.25	94.85
Ensemble (w/o simple embeddings)	97.03	93.97	96.59	95.48

4. Discussion

In this work, we considered a deep learning approach for a multiclass classification problem. In particular, we used the Pedianet database as a source of information to classify free-text diagnoses reporting no otitis, otitis (not media or acute), OM (not acute), AOM, AOM with perforation, or recurrent AOM. We trained models using five different deep learning architectures, from which a final ensemble model was developed.

Free-text information represents a valuable resource for epidemiological surveillance. Its unstructured nature, however, presents significant challenges in the extraction and interpretation of meaningful information since the manual classification of this type of information is often time- and resource-consuming, requiring extensive human effort and expertise, which can lead to inconsistencies and errors in large datasets. The advent of deep learning techniques has revolutionized this landscape. These techniques, characterized by the ability to learn hierarchical representations, have shown remarkable progresses in exploiting free text. These approaches offer several advantages, such as reducing the labor dedicated to diagnosis and real-time classification. However, this is only true if these automatic systems reach performances comparable to human levels [19,20].

The present study's results highlight the feasibility of deep learning models for automatically classifying free-text information reported in medical records, in line with previous research in pediatrics. A machine learning approach was adopted for the automatic extraction of free-text diagnoses reported in the emergency department records in Nicaragua with a median accuracy of almost 80% [21]. Another study on the Pedianet dataset employed a set of machine learning techniques (GLMNet, multinomial logistic regression (MAXENT), and the boosting approach LogitBoost) for predicting varicella cases. The performance on the test set showed that the highest predictive values were reached with the boosting algorithm (positive predictive value 63.1, 95% CI 42.7–83.5, and negative predictive value 98.8, 95% CI 98.3–99.3) [22]. In another study using the same Pedianet database, deep learning approaches were demonstrated to be feasible for detecting varicella cases for epidemiological surveillance, showing an area under the ROC curve (AUC-ROC) of 95.30% [23]. The results of this study are also consistent with a recent study on diagnoses evaluation for pediatric diseases from EHRs by Liang et al. [24] on more than one hundred million EHRs of over a million children, with excellent F1 values.

Deep learning models represent a valuable opportunity for free text exploitation since they can take advantage of datasets that are orders of magnitude larger than classical machine learning models [8]. Another advantage of deep learning approaches, particularly relevant in the context of pediatric otitis surveillance, is their ability to improve iteratively over time. This process involves starting from a model that has already undergone initial training rather than building a model from scratch for each new task. This method not only saves significant time and resources but also allows for the leveraging of the pre-existing knowledge embedded within the model. Such an approach is especially useful when considering reduced models, where some of the last layers might be excluded. By doing so, a substantial portion of the knowledge from a model that has already demonstrated high performance on a specific task can be repurposed. This existing knowledge base serves as an invaluable starting point for training a new model on a different yet related task [25]. For example, a model trained to identify patterns in general pediatric diseases can be fine-tuned to specifically recognize the nuances of otitis in children. This capability to transfer and adapt learned patterns across different but similar tasks is a hallmark of deep learning's flexibility and efficiency. It enables the development of more specialized

models that can address unique challenges in pediatric epidemiology, leading to more accurate diagnoses and better-informed healthcare strategies. This means, on the one hand, that our model could possibly be useful as a basis for training other deep learning models to classify different infections. Furthermore, considering that deep learning models can be merged to combine their knowledge, our results and methodology may also be of interest for improving other deep learning models. By integrating our model with others, the collective intelligence can address more complex or rare infections, offering a robust, comprehensive tool for medical professionals. This collaborative aspect underscores the transformative potential of deep learning in healthcare.

Recently, the advent of Generative Pre-trained Transformer (GPT) models has opened new possibilities in natural language processing, particularly in biomedical applications [26]. An increasing number of studies have presented a variety of applications using GPT models in biomedical contexts [27], with initial studies showing promising results [28,29]. GPT models offer distinct advantages over in-house-developed models; for example, they eliminate the need for text pre-processing, training, and managing computational aspects through cloud-based solutions. However, using third-party models such as GPT models raises significant bioethical [30] and privacy concerns, particularly regarding the sharing of sensitive data. In addition, the development of in-house natural language processing models allows researchers to have greater control over algorithm development and functionality, ensuring a customized approach to specific research needs. This is critical in the medical field, where accuracy and the ability to understand and modify the mechanisms underlying the model can have significant implications for both the reliability of results and patient outcomes. GPT models, while powerful in processing and generating text, often lack transparency in their operational mechanisms and decision-making processes. One major issue is the difficulty in understanding how GPT models arrive at specific outputs. In a clinical setting, it is crucial to trace the reasoning behind a diagnosis or treatment recommendation. This lack of interpretability can lead to skepticism and resistance from healthcare professionals, who are accustomed to evidence-based practices and clear, explainable decision pathways.

Exploiting Free-Text Information in Biomedical Research: Practical Implications

In the biomedical domain, the integration of underutilized free-text data with structured information holds significant potential for advancing research and clinical care [31]. Traditionally, biomedical research has heavily relied on structured data due to its ease of quantification and analysis. However, this approach often overlooks the rich, nuanced information contained in free text, such as doctors' notes, patient journals, and clinical narratives, which can provide invaluable insights into patient care and disease progression. One primary advantage of integrating these data types is the enrichment of the data pool with qualitative insights. Free-text sources frequently contain subjective patient experiences, specific symptom descriptions, and detailed treatment responses, all of which are typically absent in structured datasets. By incorporating these elements, researchers can achieve a more holistic understanding of medical conditions, leading to better-informed hypotheses and more comprehensive study designs. Additionally, free-text data can serve as a crucial tool for hypothesis generation and validation. It can reveal previously unrecognized correlations or patterns, prompting new research inquiries and supporting existing theories. For example, the detailed descriptions of patient responses to treatments in clinical notes can inform more effective therapeutic strategies. Furthermore, integrating free text with structured data can significantly enhance the accuracy of machine learning models used in biomedical research. The nuances and depth of information in free text can complement the structured data, providing a richer dataset for algorithm training and improving the predictive power of these models. In summary, leveraging often underutilized free-text data in conjunction with structured information can help to uncover new dimensions of medical research, enhance the depth and quality of analyses, and, ultimately, contribute to more effective and personalized patient care in the biomedical field.

5. Conclusions

Our analysis confirmed the potential of deep learning models in identifying and classifying diagnoses from free text. These methodologies could be adopted in other healthcare databases to improve healthcare research and limit human errors and time-speeding database interrogations.

Monitoring for otitis incidence and prevalence from pediatricians' diaries can be automated with accuracy and timeliness via deep learning models. This approach significantly enhances the efficiency and reliability of data collection in pediatric epidemiology. By harnessing the power of deep learning, the model can swiftly process vast amounts of text data, extracting relevant information with a high degree of precision. This capability ensures that the surveillance of otitis is not only faster but also more accurate, enabling healthcare professionals to identify trends and outbreaks promptly. Such timely and accurate monitoring is vital for implementing effective public health interventions and improving outcomes for pediatric patients.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jpm14010028/s1>, Table S1: Regular expressions used to filter possible cases of otitis from the Pedianet databases. The final regular expression applied was the disjunction of the reported expressions (i.e., linked with the "OR" Boolean operator). Table S2: Performances for the single kernel CNN architectures. Table S3: Performances for the sequential kernel CNN architectures. Table S4: Performances for the parallel kernel CNN architectures. Table S5: Performances for the deep-parallel kernel CNN architectures.

Author Contributions: Conceptualization, D.G. and C.L.; methodology, C.L., G.G.-H. and D.G.; software, A.M., D.W. and C.L.; formal analysis, C.L.; investigation, E.B. and D.D.; data curation, E.B. and D.D.; writing—original draft preparation, C.L.; writing—review and editing, G.L. and E.B.; visualization, G.P.; supervision, L.C., G.G.-H., C.G. and D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval was not required for the study on human participants, in accordance with the local legislation and institutional requirements. The Internal Scientific Committee approved the study and provided access to anonymized data.

Informed Consent Statement: Written informed consent from the participant's legal guardian/next of kin was not required to participate in this study in accordance with national legislation and institutional requirements.

Data Availability Statement: The data analyzed in this study are subject to the following licenses/restrictions: The data used in this study cannot be made available in the article, the Supplementary Materials, or in a public repository due to Italian data protection laws. The anonymized datasets generated during and/or analyzed during the current study will be provided upon reasonable request to the corresponding author after written approval is granted by the Internal Scientific Committee. Requests to access these datasets should be directed to the Internal Scientific Committee (info@pedianet.it).

Acknowledgments: The authors thank all the family pediatricians collaborating in Pedianet: Alongi Angelo, Angelini Roberta, Avello Giovanni, Azzoni Lucia, Balliana Franco, Barbazza Maria Carolina, Barberi Frandanisa Maria, Barbieri Patrizia, Belluzzi Gabriele, Benetti Eleonora, Bezzi Roberto, Bit Junior, Boe Franca, Bollettini Stefano, Bonfigli Emanuela, Bruna Andrea, Brusaterra Ivana, Budassi Roberto, Caccini Massimo, Cantalupi Laura, Cantarutti Luigi, Caprio Luigia, Castaldo Massimo, Castelli Stefano, Castronuovo Serenella, Cavedagni Monica, Censini Stefania, Cera Giuseppe Egidio, Ciscato Carla, Clerici Schoeller Mariangela, Collacciani Giuseppe, Comaita Fabrizio, Conte Ugo Alfredo, Costanzo Nicola, Cozzani Sandra, Cuboni Giancarlo, Curti Valentino, D'Amanti Vito Francesco, De Angelis Rita, De Clara Roberto, De Marchi Annamaria, De Nicolò Emanuele, Del Bono Gian Piero, Del Ponte Gigliola, Dell'Antonia Fabio, Di Giampietro Tiziana, Di Mauro Giuseppe, Di Renzo Anna Paola, Di Santo Giuseppe, Dolci Marco, Doria Mattia, Drago Stefano, Falco Pietro, Fama Mario, Faraci Marco, Favilli Tania, Federico Mariagrazia, Felice Michele, Ferrara Enrico, Ferrarese Marta, Ferretti Michele, Forcina Paolo, Frattini Claudio Paolo, Frison Ezio, Fusco Fabrizio, Gallo Giovanni, Galvagno

Andrea, Gentili Alberta, Gentilucci Pierfrancesco, Giampaolo Giuliana, Giancola Giuseppe, Giaretta Letizia, Girotto Silvia, Gobbi Costantino, Grelloni Mauro, Grugnetti Mirco, Lagrasta Urania Elisabetta, Landi Massimo, Lasalvia Paola, Letta Maria Rosaria, Lietti Giuseppe, Lista Cinzia, Lucantonio Riccardo, Luise Francesco, Luotti Diego, Macropodio Nadia, Marine Francesca, Mariniello Lorenzo, Marostica Gabriella, Masotti Sergio, Meneghetti Stefano, Milani Massimo, Milone Stella Vittoria, Monteleone Angela Maria, Mussinu Pierangela, Muzzolini Carmen, Nicoloso Flavia, Olimpi Laura Maria, Palma Maria Maddalena, Pandolfini Vittorio, Pasinato Angela, Passarella Andrea, Pazzola Pasquale, Perri Danilo, Pescosolido Silvana Rosa, Petrazzuoli Giovanni, Petrotto Giuseppe, Picco Patrizia, Pirola Ambrogina, Pisanello Lorena, Pittarello Daniele, Porro Elena, Profumo Elisabetta, Puma Antonino, Ragazzon Ferdinando, Rosas Paolo, Rosignoli Rino, Rossitto Mariella, Ruffato Bruno, Ruggieri Lucia, Ruscitti Annamaria, Russo Annarita, Salamone Pietro, Sambugaro Daniela, Saretta Luigi, Sarno Vittoria, Sciolla Nico Maria, Semenzato Flavio, Senesi Paolo, Silvan Carla, Spanevello Valter, Speciale Sergio Maria, Speranza Francesco, Sticco Maura, Storelli Francesco, Tamassia Gianni, Tambaro Paolo, Toffol Giacomo, Tonelli Gabriele, Tummarello Angelo, Ulliana Antonella, Venditti Sergio, Volpe Concetta, and Volpe Francescopaolo, Vozi Aldo. The authors would also like to thank the PENTA Foundation.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jamal, A.; Alsabea, A.; Tarakmeh, M.; Safar, A. Etiology, Diagnosis, Complications, and Management of Acute Otitis Media in Children. *Cureus* **2022**, *14*, e28019. [[CrossRef](#)] [[PubMed](#)]
- Barbieri, E.; Donà, D.; Cantarutti, A.; Lundin, R.; Scamarcia, A.; Corrao, G.; Cantarutti, L.; Giaquinto, C. Antibiotic Prescriptions in Acute Otitis Media and Pharyngitis in Italian Pediatric Outpatients. *Ital. J. Pediatr.* **2019**, *45*, 1–9. [[CrossRef](#)] [[PubMed](#)]
- Kaur, R.; Morris, M.; Pichichero, M.E. Epidemiology of Acute Otitis Media in the Postpneumococcal Conjugate Vaccine Era. *Pediatrics* **2017**, *140*, e20170181. [[CrossRef](#)] [[PubMed](#)]
- de Sévaux, J.L.; Venekamp, R.P.; Lutje, V.; Hak, E.; Schilder, A.G.; Sanders, E.A.; Damoiseaux, R.A. Pneumococcal Conjugate Vaccines for Preventing Acute Otitis Media in Children. *Cochrane Database Syst. Rev.* **2020**, *11*, CD001480. [[PubMed](#)]
- Monasta, L.; Ronfani, L.; Marchetti, F.; Montico, M.; Vecchi Brumatti, L.; Bavcar, A.; Grasso, D.; Barbiero, C.; Tamburlini, G. Burden of Disease Caused by Otitis Media: Systematic Review and Global Estimates. *PLoS ONE* **2012**, *7*, e36226. [[CrossRef](#)] [[PubMed](#)]
- Kim, M.O.; Coiera, E.; Magrabi, F. Problems with Health Information Technology and Their Effects on Care Delivery and Patient Outcomes: A Systematic Review. *J. Am. Med. Inform. Assoc.* **2017**, *24*, 246–250. [[CrossRef](#)] [[PubMed](#)]
- DeLisle, S.; South, B.; Anthony, J.A.; Kalp, E.; Gundlapalli, A.; Curriero, F.C.; Glass, G.E.; Samore, M.; Perl, T.M. Combining Free Text and Structured Electronic Medical Record Entries to Detect Acute Respiratory Infections. *PLoS ONE* **2010**, *5*, e13377. [[CrossRef](#)] [[PubMed](#)]
- Qian, K.; Burdick, D.; Gurajada, S.; Popa, L. Learning Explainable Entity Resolution Algorithms for Small Business Data Using SystemER. In Proceedings of the 5th Workshop on Data Science for Macro-Modeling with Financial and Economic Datasets, Amsterdam, The Netherlands, 30 June 2019; pp. 1–6.
- Chakraborty, C.; Bhattacharya, M.; Pal, S.; Lee, S.-S. From Machine Learning to Deep Learning: An Advances of the Recent Data-Driven Paradigm Shift in Medicine and Healthcare. *Curr. Res. Biotechnol.* **2023**, 100164. [[CrossRef](#)]
- Yousefinaghani, S.; Dara, R.; Poljak, Z.; Bernardo, T.M.; Sharif, S. The Assessment of Twitter’s Potential for Outbreak Detection: Avian Influenza Case Study. *Sci. Rep.* **2019**, *9*, 18147. [[CrossRef](#)]
- Pedianet. Available online: <http://Pedianet.It/En> (accessed on 15 December 2023).
- Goycoolea, M.V.; Hueb, M.M.; Ruah, C. Definitions and Terminology. *Otolaryngol. Clin. N. Am.* **1991**, *24*, 757–761. [[CrossRef](#)]
- Bojanowski, P.; Grave, E.; Joulin, A.; Mikolov, T. Enriching Word Vectors with Subword Information. *Trans. Assoc. Comput. Linguist.* **2017**, *5*, 135–146. [[CrossRef](#)]
- Lanera, C. *Development and Application of Machine Learning Techniques for Text Analyses and Classification in Clinical Research*; University of Padova: Padua, Italy, 2023.
- Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
- Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 448–456.
- Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors. *arXiv* **2012**, arXiv:1207.0580.
- Keskar, N.S.; Mudigere, D.; Nocedal, J.; Smelyanskiy, M.; Tang, P.T.P. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. *arXiv* **2016**, arXiv:1609.04836.
- Topol, E.J. High-Performance Medicine: The Convergence of Human and Artificial Intelligence. *Nat. Med.* **2019**, *25*, 44–56. [[CrossRef](#)] [[PubMed](#)]
- Keane, P.A.; Topol, E.J. With an Eye to AI and Autonomous Diagnosis. *NPJ Digit. Med.* **2018**, *1*, 40. [[CrossRef](#)]

21. Lorenzoni, G.; Bressan, S.; Lanera, C.; Azzolina, D.; Da Dalt, L.; Gregori, D. Analysis of Unstructured Text-Based Data Using Machine Learning Techniques: The Case of Pediatric Emergency Department Records in Nicaragua. *Med. Care Res. Rev.* **2021**, *78*, 138–145. [[CrossRef](#)]
22. Lanera, C.; Berchiolla, P.; Baldi, I.; Lorenzoni, G.; Tramontan, L.; Scamarcia, A.; Cantarutti, L.; Giaquinto, C.; Gregori, D. Use of Machine Learning Techniques for Case-Detection of Varicella Zoster Using Routinely Collected Textual Ambulatory Records: Pilot Observational Study. *JMIR Med. Inform.* **2020**, *8*, e14330. [[CrossRef](#)]
23. Lanera, C.; Baldi, I.; Francavilla, A.; Barbieri, E.; Tramontan, L.; Scamarcia, A.; Cantarutti, L.; Giaquinto, C.; Gregori, D. A Deep Learning Approach to Estimate the Incidence of Infectious Disease Cases for Routinely Collected Ambulatory Records: The Example of Varicella-Zoster. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5959. [[CrossRef](#)]
24. Liang, H.; Tsui, B.Y.; Ni, H.; Valentim, C.C.; Baxter, S.L.; Liu, G.; Cai, W.; Kermany, D.S.; Sun, X.; Chen, J. Evaluation and Accurate Diagnoses of Pediatric Diseases Using Artificial Intelligence. *Nat. Med.* **2019**, *25*, 433–438. [[CrossRef](#)]
25. Dahl, G.E.; Yu, D.; Deng, L.; Acero, A. Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition. *IEEE Trans. Audio Speech Lang. Process.* **2011**, *20*, 30–42. [[CrossRef](#)]
26. Xue, V.W.; Lei, P.; Cho, W.C. The Potential Impact of ChatGPT in Clinical and Translational Medicine. *Clin. Transl. Med.* **2023**, *13*. [[CrossRef](#)] [[PubMed](#)]
27. Zhang, P.; Kamel Boulos, M.N. Generative AI in Medicine and Healthcare: Promises, Opportunities and Challenges. *Future Internet* **2023**, *15*, 286. [[CrossRef](#)]
28. Takagi, S.; Watari, T.; Erabi, A.; Sakaguchi, K. Performance of GPT-3.5 and GPT-4 on the Japanese Medical Licensing Examination: Comparison Study. *JMIR Med. Educ.* **2023**, *9*, e48002. [[CrossRef](#)]
29. Brin, D.; Sorin, V.; Vaid, A.; Soroush, A.; Glicksberg, B.S.; Charney, A.W.; Nadkarni, G.; Klang, E. Comparing ChatGPT and GPT-4 Performance in USMLE Soft Skill Assessments. *Sci. Rep.* **2023**, *13*, 16492. [[CrossRef](#)]
30. Li, H.; Moon, J.T.; Purkayastha, S.; Celi, L.A.; Trivedi, H.; Gichoya, J.W. Ethics of Large Language Models in Medicine and Medical Research. *Lancet Digit. Health* **2023**, *5*, e333–e335. [[CrossRef](#)]
31. Scheurwegs, E.; Luyckx, K.; Luyten, L.; Daelemans, W.; Van den Bulcke, T. Data Integration of Structured and Unstructured Sources for Assigning Clinical Codes to Patient Stays. *J. Am. Med. Inform. Assoc.* **2016**, *23*, e11–e19. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.