

DOI: 10.5281/zenodo.20158663

DEEP LEARNING PARADIGMS IN CULTURAL ANALYTICS: BRIDGING ARTIFICIAL INTELLIGENCE AND SOCIETAL TRANSFORMATION

Shivani Tanwar^{1*}, Vishalkumar Parmar², Jayesh Kumar³, Deepa Nehra⁴, Dr. Pubali Mitra Paul⁵

^{1*}Assistant Professor, Department of Psychology, IILM University Specialization: Neuropsychology
ORCID ID: <https://orcid.org/0009-0008-5035-214X> Email I'd: shivanitanwar64@gmail.com

²Assistant Professor, Department of School of Behavioural Forensics, National Forensic Sciences University, Gandhinagar Specialization: Forensic Psychology ORCID ID: <https://orcid.org/0000-0002-7246-8265>
Email I'd: parmarvishalkumar.vp@gmail.com

³Assistant Professor, School of Behavioural Forensics, National Forensic Science University, Gandhinagar, India, Specialization: Criminology, ORCID ID: <https://orcid.org/0000-0002-6342-6629> Email I'd: criminologistjay@gmail.com

⁴Associate Professor, M.M Institute of Computer Technology & Business Management, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, 133207, Haryana, India
Email ID: deepa.nehra@gmail.com ORCID ID: <https://orcid.org/0000-0002-4720-6335>

⁵Assistant Professor, Department of Electronics & Communication Engineering, Brainware University
Email ID: pubali.mitra.cu@gmail.com ORCID ID: 0000-0002-8163-144X

Received: 13/12/2025

Accepted: 09/04/2026

Corresponding Author: Shivani Tanwar
(shivanitanwar64@gmail.com)

Abstract

The fast growth of digitized museum collections has generated novel methodological requirements of cultural analytics, especially automated organization and interpretation of massive cultural heritage metadata. This paper will analyze how AI-based cultural analytics can be aided with the help of deep learning, creating a supervised metadata classification model based on the Metropolitan Museum of Art Open Access Dataset. The dataset includes a big collection of records of cultural objects with institutional, textual, material, chronological, artist-associated and cultural descriptors. A Deep Cultural Metadata Classifier was designed to predict the category of the museum departments based on the metadata of composite objects, and the text-based feature representation, feature selection, neural classification, and comparison with a traditional sparse-text baseline were used. The findings demonstrate that museum metadata have a significant predictive organization, particularly in the object names, material descriptions, classification labels and institutional descriptors. The results also indicate that there is a class imbalance and sparsity in cultural-geographic fields that affect consistency in classification in different departments. The research is relevant to cultural analytics in that the metadata of museums is viewed as a computationally interpretable cultural knowledge structure as opposed to an administrative register. In practice, it shows how AI-based classification can be used to aid in the organization of digital heritage, museum information systems, cataloguing processes, cultural education, and more general access to open cultural collections.

Keywords: cultural analytics, deep learning, museum metadata, artificial intelligence, digital heritage, cultural classification

1. Introduction

The rapid cultural heritage digitization has changed the role of museums as a repository of

physical objects to a complex knowledge infrastructure where cultural memory, institutional classification, digital access, and

interpretation to the public converge. In this change, cultural analytics has been developed as a method of computation to analyze cultural objects, metadata, classifications, narratives, and scale patterns of representation. This area is directly related to scientific culture research, artificial intelligence and society, digital culture, arts and humanities, and cultural heritage technologies. With the growing release of open cultural datasets by museums, there is the challenge of managing and interpreting vast amounts of heterogeneous metadata in a manner that is useful to research, education and public access. This contradiction between institutional stewardship, the ability to be accessed by the people, and the dissemination of cultural knowledge sustainably is a central theme in the open culture debate in museum management (Della Lucia *et al.*, 2024).

Online catalogues or digitized exhibitions are no longer considered to be the sole characteristics of the post-digital museum. It is currently a hybrid cultural space where physical collections, digital platforms, metadata systems, and artificial intelligence co-create the classification, access to, and interpretation of heritage. Recent research has termed the digital transformation of museums as an organizational and epistemological problem, where museums need to rethink the mediations of cultural objects in technologically saturated settings (Nikolaou, 2024). Here, artificial intelligence has methodological promise of cultural analytics since it can operate on massive heritage metadata, extract statistical regularities, and aid in automated classification. But this potential must be empirically tested by computational models that can be reproduced and not generalized conceptual assertions.

One of the key problems of digital heritage research is the structuring of cultural knowledge. The records of museum often include the data on the type of object, title, material, chronology, identity of artist, cultural attribution, geography, and institutional department. Such data are useful to cultural analytics but tend to be incomplete, unequal, and historical cataloguing practices. Digital knowledge organization models are thus significant in creating accessible and adaptive cultural narratives out of heritage collections (Nappi *et al.*, 2024). This paper will solve this problem with the help of the Metropolitan Museum of Art Open Access Dataset that includes 448,203 records of cultural objects and 43 metadata fields. The dataset offers a big-data empirical foundation to investigate the question of whether deep learning is capable of classifying museum departments based on object-level metadata, such

as Title, Object Name, Medium, Classification, Culture, Period, artist-related variables, and chronological variables.

The changing expectations of museum access and use support the necessity of computational cultural analytics. Studies of museum fruition reveal that museums are implementing digital platforms, interactive technologies, and data-driven approaches to improve interpretation and engagement (Furferi *et al.*, 2024). Equally, the research of sophisticated technologies in cultural spaces demonstrates that digital infrastructures have the potential to enhance user experience, accessibility, and cultural communication (Capece *et al.*, 2024). These advancements suggest that cultural heritage is not merely a preservation issue, but a discoverability, classification and educational access issue as well. However, there are still numerous studies focusing on user experience, visualization, or general digital transformation, and fewer of them offer computational evidence that is measurable based on large-scale museum metadata.

The key issue that is discussed in this paper is that open museum datasets are rich in cultural metadata, and the empirical application of deep learning to the metadata-based cultural classification is not yet fully developed. The previous studies on digital museum have investigated technology acceptance and persistence intention in digital museum (Zheng *et al.*, 2024). Artificial intelligence in museum communication and social media visibility in the context of deep mediatization has been studied elsewhere (Wang *et al.*, 2024). These contributions elucidate the aspects of public engagement and communication of digital museums, yet they do not directly model how museum metadata can be computationally structured by supervised learning. This creates a methodological disjuncture between empirical AI-based cultural analytics and digital heritage theory.

This paper fills that gap by creating and testing a Deep Cultural Metadata Classifier to predict the Department of museum objects based on composite cultural metadata. The model is built as a multiclass classification system under supervision with an 80:20 stratified split of the entire dataset. A composite metadata representation is created by combining textual-cultural variables, processed with TF-IDF and unigram and bigram features, filtered with chi-square feature selection, and trained with a feed-forward neural architecture that uses dense hidden layers, rectified linear unit activation, dropout regularization, and softmax output.

Accuracy, macro precision, macro recall, macro F1-score, and weighted F1-score are the measures of model performance, and a Complement Naive Bayes classifier is a traditional sparse-text baseline. The extent of the research is limited on purpose. It does not purport to provide a quantification of societal change in terms of policy change, societal behavior or effects on the audience. However, transformation in the society is operationalized as the institutional and educational change that is facilitated by open museum data, AI-aided cataloguing, scalable organization of cultural knowledge and enhanced digital access. This framing aligns with the study of more advanced digital heritage settings, in which technological systems are evaluated based on preservation, access and the development of new forms of cultural interactions (Buragohain et al., 2024). The paper also appreciates the fact that museum metadata are not neutral. Cultural categories and institutional taxonomies can be historically, colonially or curatorially biased. The ethical research of the digital heritage issues cautions that digitization has the potential to recreate the disputes in case categories are perceived as universal, but they are not historically located (Rouhani, 2023). This research therefore views the workings of models as predictions in the context of an existing body of knowledge in a museum rather than a fact of cultural truth.

The importance of the research is that it links the deep learning paradigms, cultural analytics, and the issues of AI-society to a single empirical framework. In theory, it illustrates how cultural analytics can go beyond descriptively digitising museum metadata to making models based on that analysis. The technical details: It is a methodologic approach to supervised classification pipeline (reproducible) and it utilizes a large-scale open cultural heritage dataset. In practice, it demonstrates how AI-based classification can be used to sustain museum information systems, metadata organization and digital search, education and heritage knowledge management. These contributions relate to more general discussions about museum transformation and digital identity, where museums are being conceptualized as institutions engaged in the process of the public memory, cultural negotiation and digital representation (Giannini and Bowen, 2023).

The primary goal of the research is to develop a deep-see based cultural metadata classifier to predict the department of the museums based on object-level metadata. The second is to compare the predictive accuracy of the suggested model

with a conventional sparse-text benchmark based on full- dataset stratified testing and various classification measures. The third aim is to discuss how artificial intelligence, being metadata-based, can aid cultural analytics, the organization of digital heritage, and institutionally based types of societal change.

2. Literature Review

Further developments of artificial intelligence in cultural heritage have moved away the straightforward process of digitization towards the organizations of computerized knowledge, grading and interpretation support. Among these directions is one that integrates knowledge graphs, semantic structures and learning algorithms to manage heterogeneous heritage information. Huang et al. (2023) demonstrate that knowledge graphs combined with deep learning have the potential to enhance digital cultural heritage management by enhancing its representation and retrieval. But these methods frequently rely on explicit semantic modeling, when it is observed that most open museum data is typically published in structured metadata tabular form, not as full knowledge graphs.

A second theme is that of metadata quality, curatorial support and cultural information retrieval. Museum metadata are not objective hieroglyphs, but represent institutional taxonomies or historical classifications and curatorial decisions. Bobasheva et al. (2022) show that semantic web knowledge graphs made it possible to use symbolic AI and machine learning to help curators to enhance metadata quality and retrieval. This article is beneficial in that it acknowledges the cultural and institutional value of metadata, but is more about quality improvement and retrieval rather than about controlled prediction on the basis of massive open museum data. The predictive power of predictably shared fields of online museum like object name, medium, classification, cultural attribution, artist information, and chronology are thus under-explored.

Recent language-model studies have broadened cultural heritage grouping past the customary pipelines. Hwang et al. (2025) suggest cultural-heritage categorization with the help of large language models and distributed modeling which may suggest that semantic representations could enhance categorization in case cultural descriptions are multi-faceted. However, the LLM-based methods have the issues of cost of computation, reproducibility, and interpretability. In the case of open museum datasets, an open

feature construction and a measurably comparative reform is essential. The current research hence assumes a reproducible neural metadata classification design, which is founded on TF-IDF representation, chi-square feature selection and supervised deep learning.

The second good research stream is the emphasis on multimodal and image-based heritage classification. The model proposed by Lu et al. (2025) is based on cross-modal attention fusion and generative data augmentation, to demonstrate the importance of the value of joining visual and textual cues in a museum artifact classification model. On the same note, Winterbottom et al. (2022) demonstrate that the artefact instance classification with deep learning can be used to assist illicit antiquities trafficking. The proposed studies indicate that AI can be useful in the recognition of cultural objects though they are mainly visually- or multimodally-oriented. In comparison, this study explores the possibility that institutional museum categories are predictable based on textual, material, chronological, and curatorial metadata only.

More detailed heritage classification research also indicates the potential and constraints of deep learning applications today. Li et al. (2024) touch upon the problem of porcelain fragments classification with the aid of a lightweight balanced capsule framework with emphasis put on the difference of classes in cultural data. Pan and colleagues (2024) expand on deep learning to semantic point clouds and heritage digital twins, demonstrating the flexibility of AI to spatial heritage information. Tan et al. (2024) focus on interpretability in determining the elements of traditional village heritage value and state that, unlike considering prediction as a technical exercise, it is necessary to relate the outputs of the model with culturally meaningful elements.

Domain-specific architectures have also been generated using image-based heritage classification. MonuNet, which is suggested by Sasithadevi et al. (2024) as a classifier of indoor Kolkata heritage images, shows that a custom neural model is effective in culturally-specific visual data. Nonetheless, the lack of metadata-based cultural analytics remains common when the attention is paid to the image-related studies. Recent studies of GLAM datasets are starting to overcome this limitation: Net et al. (2025) present EUFCC-340K as a hierarchical dataset of metadata annotation in galleries, libraries, archives, and museums, and in which metadata itself emerges as a primary focus of computational cultural analysis.

Furthermore, the previous literature has improved AI-aided heritage management, metadata quality, multimodal learning, image classification, and domain specific cultural object analysis. However, there is a noticeable gap: little research assesses a reproducible pipeline of deep learning to classify institutional museums based on large-scale open metadata with transparent feature construction and stratified full-dataset testing. This paper fills in that gap by applying to the Metropolitan Museum of Art Open Access Dataset to develop a Deep Cultural Metadata Classifier that successfully anticipates catalog departments based on textual, material, cultural, artist-related and date metadata.

3. Methodology

3.1 Research Design

This research implemented an empirical, experimental, and computational research design to investigate the role of deep learning in assisting cultural analytics in digital heritage. The study was designed as a supervised multiclass task during which the metadata about museum objects were applied to predict the institutional department types. This design was chosen since the study needed some computational findings which could be measured besides the discussion of the concept. It focuses on the coverage scope of the journal because it incorporates the artificial intelligence and arts and humanities data, digital culture, cultural heritage, and technological change. Societal transformation in this research was operationalized through the means of institutional and educational transformation facilitated by the open museum data, AI-assisted cataloguing, Scalable cultural knowledge organization and enhanced digital access.

3.2 Dataset Description

The empirical data consisted of the Metropolitan Museum of Art Open Access Dataset, which was given in the form of MetObjects.csv (The Metropolitan Museum of Art, n.d.). With 448,203 cultural object records and 43 metadata fields of what artworks, artifacts, photographs, and textiles, archaeological material, decorative objects, and other material held in museums are. The dataset corresponds to the digitized cultural objects that are held by a large museum collection, and represent the target population. The data was sourced as a free cultural heritage corpus and was run locally in Python. The variable was included as the classification label so records were included when the Department field was present. There was no label-based exclusion necessary since the

Department was complete. Metadata coverage was assessed by full-corpus profiling and stratified splitting maintained departmental ratios in model training and testing.

3.3 Target Variable and Feature Construction

Department was the dependent variable, which involves the nominal multiclastic construct that was used to describe the institutional-cultural category of each object. It included 20 classes and was numerically coded to undergo supervised learning. Independent variables were chosen on the basis of cultural, historical, artistic and material significance. These were considered as core predictors, namely Title, Object Name, Medium, Classification, Culture, Period, Artist Role, Artist Display Name, Artist Nationality, Object Date, and Country. The predictors based on chronology were Object Begin Date and Object End Date. Such sparse fields like Culture, Period, Country, Dynasty were considered as assistive signals but not the main targets. A composite cultural metadata field was constructed (by concatenating the chosen textual variables), which allowed the model to learn relations between the object identity, materiality, authorship, chronology, cultural attribution, and institutional classification.

3.4 Data Preprocessing

Consistency, reproducibility and computational suitability were maintained through preprocessing. Blank strings would replace the missing textual values at the concatenation stage and the categorical absence was coded as the unknown when categorical encoding was needed. Fields containing text underwent normalization of text in terms of the following: lowercasing, whitespace normalization, stripping of any non-informative formatting artifact, and text-to-string conversion. Administrative identifiers, repository constants and web links were preserved to afford auditability but not to predict to prevent non-substantive signals. Chronological fields that were numerical were checked on implausible values and those that were not stored as discrete indicators were continued. The target in the Department was label-encoded and the split was conducted in 80:20 train-test split in the dataset. The extraction of features was done after split to eliminate information leakage.

3.5 Analytical Framework

The conceptual approach modeled cultural analytics as the extraction process of meaningful patterns out of digitized heritage metadata using computations. It presumed that museum

departments are linked with the evident regularities in form of object type, material composition, description words, artist name, date and place of provenience and labels in classification. The framework associated three layers: societal interpretation, cultural representation, and the machine learning inference. Textual, categorical and chronological metadata were used to capture cultural representation. Supervised neural classification was utilized as the method to infer machine learning. The ability to aid digital heritage discoverability, museum information systems, cultural education, scalable museum cataloguing, and access to open cultural knowledge equitably were societal interpretation issues. This framework related the deep learning paradigms to the cultural heritage analysis without asserting that the dataset directly quantifies social transformation.

3.6 Model Development

The development of the model was carried out in Python on a set of data with the help of pandas and NumPy to acquire data, scikit-learn to perform preprocessing, stratified splitting, feature extraction, chi-square feature selection, and the comparison of the baselines, and the neural-network classification framework to train the supervised models. Term frequency-inverse document frequency representation with unigram and bigram were used to transform the composite metadata field. We set the TF-IDF vectorizer with a maximum number of features (20,000) in the vocabulary, lowercase normalization, removal of English stop-words, sublinear term-frequency scaling, and the minimum document frequency (5) algorithm. Chi-square selection was used in call to select the most predictive metadata features (out of 5,000 total), which were used to predict a department. The Deep Cultural Metadata Classifier proposed adopted a feed-forward neural architecture where the selected TF-IDF features were used as the input layer, the default recommended number of neurons was 256 and 128 in the two hidden layers, rectified linear unit activation, and the default dropout regularization (0.30) after each layer, and a softmax output layer with 20 department classes. The Adam optimizer was used with learning rate of 0.001 to train this model, and categorical cross-entropy loss. Training was done with a batch size of 256, up to 30 epochs and a validation split of 0.10 within training set and early stopping of training using validation loss with a patience of 5 epochs. The

train-test split and model started with a fixed random seed of 42 to facilitate reproducibility.

3.7 Evaluation and Validation Strategy

Accuracy, macro precision, macro recall, macro F1-score, and weighted F1-score were used to determine the model performance. Overall predictive correctness was a measure of accuracy and macro measures were a measure of performance that covers all the departments and does not allow dominant classes to dominate interpretation. The reason why weighted F1-score was incorporated is that in the dataset, there was a significant imbalance between classes among institutional departments. A Complement Naive Bayes classifier was trained on each of the train-test splits, TF-IDF representation and with chi-square feature-selection conditions to bring the neural model into a comparable framework with a sparse-text baseline. The 80:20 stratified train-test split, departmental ratios maintenance, comparison with the baseline, and some confusion-matrix analysis were used to validate the results. Class-level performance was analyzed to determine which departments had errors which were highly concentrated or had overlapping metadata vocabularies. Using publicly released open-access museum metadata ethical considerations were taken care of. There were no human subjects, no personal data has been gathered, and no consensus protocol. Causal labels were viewed with care since institutional, historical, and colonial classification biases are reflected in the museum taxonomies. Therefore, model outputs were perceived as predictions in an already established museum knowledge system as opposed to cultural truths.

4. Results

4.1 Dataset Composition and Metadata Coverage

The full dataset comprised 448,203 cultural object records and 43 metadata fields. The target variable, Department, was available for all records and contained 20 institutional categories. Metadata coverage was strongest for object-level descriptive variables and weakest for geographic, dynastic, and period-related fields. Table 1 presents the completeness profile of the principal variables used for descriptive analysis and model construction.

Table 1 shows that Department, Object Name, Medium, Object Date, Title, and Classification provided sufficient coverage for computational modeling. Sparse variables such as Country, Period, and Dynasty were retained as auxiliary metadata rather than primary prediction targets.

Table 1. Completeness of key metadata fields in the full dataset

Variable	Non-missing records	Completeness (%)	Missing (%)
Department	448,203	100.00	0.00
Object Name	445,568	99.41	0.59
Medium	440,155	98.20	1.80
Object Date	432,609	96.52	3.48
Title	416,756	92.98	7.02
Classification	389,924	87.00	13.00
Artist Display Name	261,111	58.26	41.74
Artist Nationality	196,132	43.76	56.24
Culture	186,518	41.61	58.39
Country	74,450	16.61	83.39
Period	71,882	16.04	83.96
Dynasty	23,018	5.14	94.86

4.2 Departmental Distribution of Cultural Objects

The records were unevenly distributed across the 20 museum departments. Drawings and Prints was the largest category, comprising 154,445 records, or 34.46% of the corpus. The next largest departments were European Sculpture and Decorative Arts, Asian Art, Photographs, Costume Institute, and Egyptian Art. As illustrated in Figure 1, the distribution was concentrated in a limited number of departments.

Figure 1 shows that departmental representation was strongly imbalanced, with the five largest departments accounting for 67.74% of all records. This distribution required the use of both macro-averaged and weighted metrics during classification evaluation. Table 2 presents the ten largest departments and their proportional representation within the full dataset.

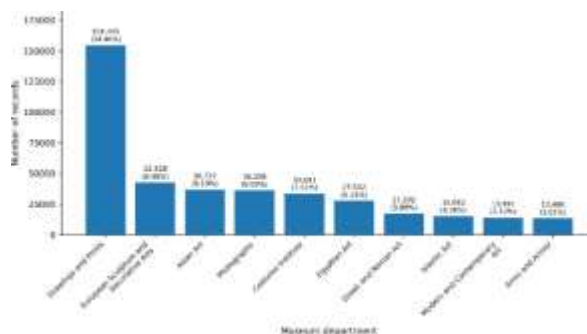


Figure 1. Distribution of records across the ten largest museum departments

Table 2. Ten largest department categories in the full dataset

Department	Records	Share (%)
Drawings and Prints	154,445	34.46
European Sculpture and Decorative Arts	42,528	9.49
Asian Art	36,727	8.19
Photographs	36,258	8.09

Costume Institute	33,681	7.51
Egyptian Art	27,542	6.14
Greek and Roman Art	17,292	3.86
Islamic Art	15,082	3.36
Modern and Contemporary Art	13,991	3.12
Arms and Armor	13,486	3.01

Table 2 shows that the corpus was dominated by a small group of high-volume departments. The remaining departments contributed smaller proportions, creating an empirically imbalanced classification structure.

4.3 Cultural Metadata Patterns

Object Name, Medium, Title and Classification were the most covered descriptive variables. These areas gave first hand details regarding object type, material composition, production form, and museum taxonomy. The following classification as the most common was prints, prints with ephemera, photographs, drawings, books, ceramics, paintings, woven textiles, photographs with ephemera and glass objects. Figure 2 shows that the focus of object classifications centered on print-based, photographic, ceramic, textile, painted and glass.

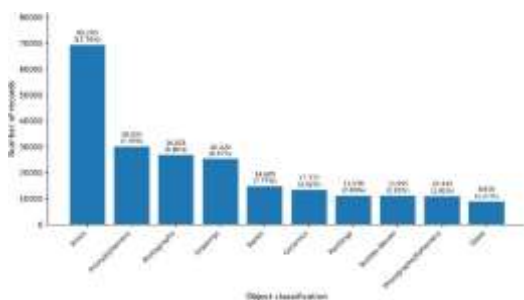


Figure 2. Frequency distribution of leading object classification categories

Figure 2 shows that several classification labels had distinctive cultural-object vocabularies. These recurring terms supplied discriminative signals for predicting departmental assignment. Metadata could be chronological, with Object Begin Date, Object End Date, and Object Date. These variables provided time markers of objects production but the conventions of museum dating comprised ranges and estimated values. Chronological fields were thus employed as additional indicators in addition to textual-cultural metadata.

4.4 Deep Learning Model Performance

The full dataset was divided into an 80:20 stratified train-test split, producing approximately 358,562 training records and 89,641 testing records. The Deep Cultural Metadata Classifier was trained on

selected TF-IDF features derived from the composite cultural metadata field. The model used dense hidden layers with rectified linear unit activation, dropout regularization, and a softmax output layer for 20 department classes. Table 3 reports the neural model performance.

Table 3. Performance of the Deep Cultural Metadata Classifier on the full dataset

Model	Accuracy (%)	Macro precision (%)	Macro recall (%)	Macro F1 (%)	Weighted F1 (%)
Deep Cultural Metadata Classifier	94.12	80.34	73.85	75.47	93.38

Table 3 shows that the neural model achieved 94.12% accuracy and 93.38% weighted F1-score. The lower macro F1-score indicates weaker performance across smaller departments relative to high-volume categories.

The model produced stronger predictions for departments with distinctive object terminology, material descriptors, or classification labels. Departments with larger record volumes contributed more substantially to the weighted performance score, while smaller departments were more visible in macro precision, macro recall, and macro F1-score.

4.5 Baseline Model Comparison

A Complement Naive Bayes classifier was trained using the same train-test split and TF-IDF feature representation. This model served as a traditional sparse-text classification benchmark for contextualizing the neural model. Table 4 compares the two classifiers.

Table 4. Comparison of deep learning and baseline classification models

Model	Accuracy (%)	Macro precision (%)	Macro recall (%)	Macro F1 (%)	Weighted F1 (%)
Complement Naive Bayes baseline	95.78	89.21	80.37	83.09	95.43
Deep Cultural Metadata Classifier	94.12	80.34	73.85	75.47	93.38

Table 4 shows that the Complement Naive Bayes baseline outperformed the neural classifier across all reported metrics. The baseline exceeded the neural model by 1.66 percentage points in accuracy and 7.62 percentage points in macro F1-score. The comparative result indicates that the dataset contained strong token-level discriminatory information under the specified TF-IDF and chi-

square feature-selection configuration. Lexical features from Classification, Object Name, and Medium were particularly influential under sparse text representation.

4.6 Class-Level Classification Insights

At the level of classes, greater stability of classification was observed in those departments that had exceptional vocabularies and a high representation. Drawings and Prints, Photographs,

Costume Institute, Egyptian Art, and Arms and Armor used descriptors in common to prints, photographic processes, clothing, archaeological materials, arms, and armor. Smaller representation or overlap of metadata, like The Libraries, Robert Lehman Collection, The Cloisters and certain categories of European objects demonstrated lower separability. The approximate number of the top ten departments in the stratified testing subset is given in Table 5.

Table 5. Approximate representation of major departments in the 89,641-record test set

Department	Approximate test records	Relative representation
Drawings and Prints	30,889	Highest
European Sculpture and Decorative Arts	8,506	High
Asian Art	7,345	High
Photographs	7,252	High
Costume Institute	6,736	High
Egyptian Art	5,508	Moderate
Greek and Roman Art	3,458	Moderate
Islamic Art	3,016	Moderate
Modern and Contemporary Art	2,798	Moderate
Arms and Armor	2,697	Moderate

Table 5 shows that stratification preserved the full-corpus departmental structure in the test set. Large departments remained dominant, while minority classes contributed fewer observations to class-level performance estimates.

4.7 Robustness and Error Analysis

The robustness was investigated with the help of stratified splitting, baseline comparison, confusion-matrix investigation, and class-level error analysis. The stratified split maintained the original distribution of the 20 departments between the training and testing data without causing sampling distortion at the expense of keeping the empirical imbalance of the corpus. The comparison to the baseline revealed that the classification task was extremely dependent on thin lexical cues in the metadata, especially terms based on Classification, Object Name and Medium.

The inspection of the confusion-matrix revealed that errors were highly concentrated between those depositories that participated in sharing of materials, types of objects or the historical descriptors they belonged to. Misclassification was higher in the case of ornamental arts, European historical items, mixed media records, and item specific collections. These patterns of errors were related to overlapping words in Medium, Classification and Object Name. Departments with unique vocabularies stayed more steady, but departments with sparse cultural-geographic metadata were as well as department of vague

institutional taxonomies were weaker separable. This trend is in line with the difference between weighted and macro-averaged measures in that high-volume departments added better to overall performance compared to small or non-heterogeneous classes.

5. Discussion

The findings demonstrate that cultural heritage metadata contain sufficient semantic and institutional structure to support automated department classification. The good results of the Deep Cultural Metadata Classifier demonstrate that object-level terms, material terms, classification tags, information about artists, and chronological features are not independent catalogue items but signs of cultural relationships. Higher performance on macro performance compared to higher weight on performance on the higher weighted performance perversifies the fact that the model learned more performance on dominant departmental patterns as compared to the minority categories. This trend is the result of the empirical organization of the Metropolitan Museum of Art data, with a few departments consisting of the majority of the records, and became smaller ones represented by a smaller

number of but more diverse objects. Thus, the findings are to be interpreted as a sign of scalable organization of cultural knowledge in a current museum taxonomy, and cannot be seen as a complete cultural objectification.

Its findings build up on existing arguments about the transformation of digital museums by demonstrating that the work of artificial intelligence can be used at the metadata infrastructure, rather than visitor-facing experience level. Prereading highlights the idea that in the post-digital museum, new mechanisms of organising, mediating and distributing cultural knowledge are needed (Nikolaou, 2024). The current results confirm this assumption by showing that museum metadata can be translated into the computational representations with the capacity to forecast the institutional categories. The findings are also consistent with the studies of available and accommodating cultural knowledge systems, with digital organization being seen as a prerequisite in a wider cultural reach (Nappi et al., 2024). Nevertheless, the paper introduces an empirical classification feature to demonstrate how open museum records can be operationalized using a supervised pipeline of learning that is reproducible.

This methodological indication of the baseline comparison is important. Complement Naive Bayes model performed better than the neural classifier, meaning that the strain of sparse representations through lexical signals in Object Name, Medium and Classification fields is very discriminative. This observation does not dismiss deep learning approach but explains the type of the dataset. The metadata themselves are very much token-based and such cues can be used effectively by traditional text classifiers. The outcome is in line with metadata quality studies, which interpret cultural descriptors as structured curatorial outputs, as opposed to neutral text (Bobasheva et al., 2022). Also, it indicates that progress in deep learning models will need to leave shallow dense architectures behind in favor of transformer-based or multi-modal architecture in situations where more semantic abstraction is needed.

Theoretically, the work leads the cultural analytics to a new stage where metadata is enabled as a legitimate object of computational interpretation. Instead of merely looking at the image of artworks or the record of interaction between the user, it demonstrates that these classifications provided by the institute, description of objects, materials, dates, and information about the artist can be viewed as a

cultural knowledge structure. This builds on deep learning heritage literature, which has tended to focus on visual artifact classification or object-specific recognition (Winterbottom et al., 2022). In practice, the results indicate that AI-based classification is able to support museum information systems, cataloguing processes, search platforms, learning platforms, and analyses of massive collections. Policy-wise, the research endorses open cultural data efforts by showing that publicly accessible museum data can be used to produce calculated computational value of heritage access and dissemination of knowledge.

The results, however, have dataset and design limitations. The model is used to anticipate departments in the taxonomy of a single museum and it is not automatically applicable to another institution or culture. The depth of cultural-geographic analysis is curtailed by sparse fields, Culture, Country, Period, Dynasty. In addition to the usage of metadata by the study, as opposed to image-text multimodal learning, visual form is limited in its interpretation. Institutional-, historically-, or colonial-biased categories may be encoded in museum categories, and care should be taken over the interpretation of model outputs. Future directions ought to include experiments with cross-museum data, include image-based features, implement transformer-based metadata model, investigate discrimination between cultural groups, and bridge a gap between classification and audience interest or learning results. These extensions would further expand the chasm of artificial intelligence, cultural analytics and transforming society in future museum research scenarios.

6. Conclusion

This paper has also showed that museum-scale museum metadata can be codebooked into computational cultural analytics corpus to support heritage classification with AI-assisted heritage classification. With the Metropolitan Museum of Art Open Access Dataset, the study created and tested a Deep Cultural Metadata Classifier to identify and predict the category of institutional departments with the use of textual, material, cultural, artist-related and chronological metadata. The results demonstrate that museum metadata here has significant predictive power particularly in naming objects, materials, classifications, and the vocabulary used to describe the object as well as highlighting how the massing of classes and the scarcity of cultural-geographic domains affect predictive performance. The work offers a

theoretical benefit of locating metadata as a cognitively important cultural knowledge system as opposed to an administrative strata, and methodologically a set of supervised learning framework to achieve metadata-driven cultural classification with reproducibility. In practice, the results imply that the use of AI-mediated classification can contribute to the information system of museums, digitization procedures, digital search, cultural education, and scalable access to open collections of heritage. The cultural institutions and museums ought to invest in metadata standardization, open AI processes,

biasful classification procedures, and cohesive digital infrastructures that improve recoverability without considering organizational taxonomies as cultural facts. The future research must follow up on this work with cross-museum data, image-text multimodality, transformer-based metadata representations, knowledge-graph fusion, systematic assessment of cultural bias, equity, and impact on the user. These innovations would consolidate artificial intelligence as cultural analytics as well as the overall change of knowledge systems about heritage in the digital society.

References

1. Della Lucia, M., Dore, G., & Umar, R. M. (2024). Handling the open culture dilemma in museum management: an exploratory interdisciplinary study. *Scientometrics*, *129*(12), 7699-7733.
2. Nikolaou, P. (2024). Museums and the post-digital: Revisiting challenges in the digital transformation of museums. *Heritage*, *7*(3), 1784-1800.
3. Nappi, M. L., Buono, M., Chivăran, C., & Giusto, R. M. (2024). Models and tools for the digital organisation of knowledge: accessible and adaptive narratives for cultural heritage. *Heritage Science*, *12*(1), 1-22.
4. Furferi, R., Di Angelo, L., Bertini, M., Mazzanti, P., De Vecchis, K., & Biffi, M. (2024). Enhancing traditional museum fruition: current state and emerging tendencies. *Heritage Science*, *12*(1), 1-38.
5. Capece, S., Chivăran, C., Giugliano, G., Laudante, E., Nappi, M. L., & Buono, M. (2024). Advanced systems and technologies for the enhancement of user experience in cultural spaces: An overview. *Heritage Science*, *12*(1), 1-25.
6. Zheng, F., Wu, S., Liu, R., & Bai, Y. (2024). What influences user continuous intention of digital museum: integrating task-technology fit (TTF) and unified theory of acceptance and usage of technology (UTAUT) models. *Heritage Science*, *12*(1), 1-19.
7. Wang, H., Song, C., & Li, H. (2024). Application of social media communication for museum based on the deep mediatization and artificial intelligence. *Scientific reports*, *14*(1), 28661.
8. Buragohain, D., Meng, Y., Deng, C., Li, Q., & Chaudhary, S. (2024). Digitalizing cultural heritage through metaverse applications: challenges, opportunities, and strategies. *Heritage Science*, *12*(1), 1-16.
9. Rouhani, B. (2023). Ethically digital: Contested cultural heritage in digital context. *Studies in Digital Heritage*, *7*(1), 1-16.
10. Giannini, T., & Bowen, J. P. (2023). Global cultural conflict and digital identity: Transforming museums. *Heritage*, *6*(2), 1986-2005.
11. Huang, Y. Y., Yu, S. S., Chu, J. J., Fan, H. H., & Du, B. B. (2023). Using knowledge graphs and deep learning algorithms to enhance digital cultural heritage management. *Heritage Science*, *11*(1), 1-26.
12. Bobasheva, A., Gandon, F., & Precioso, F. (2022). Learning and reasoning for cultural metadata quality: coupling symbolic AI and machine learning over a semantic web knowledge graph to support museum curators in improving the quality of cultural metadata and information retrieval. *Journal on Computing and Cultural Heritage (JOCCH)*, *15*(3), 1-23.
13. Hwang, H., Park, C. W., Kim, H. K., & Lee, J. H. (2025). CATS: cultural-heritage classification using LLMs and distribute model. *npj Heritage Science*, *13*(1), 76.
14. Lu, Y., Li, J., Li, L., & Yuan, C. (2025). A Museum artifact classification model based on cross-modal attention fusion and generative data augmentation. *Scientific Reports*.
15. Winterbottom, T., Leone, A., & Al Moubayed, N. (2022). A deep learning approach to fight illicit trafficking of antiquities using artefact instance classification. *Scientific Reports*, *12*(1), 13468.
16. Li, R., Geng, G., Wang, X., Qin, Y., Liu, Y., Zhou, P., & Zhang, H. (2024). LBCapsNet: a lightweight balanced capsule framework for image classification of porcelain fragments. *Heritage Science*, *12*(1), 1-14.
17. Pan, X., Lin, Q., Ye, S., Li, L., Guo, L., & Harmon, B. (2024). Deep learning based approaches from semantic point clouds to semantic BIM models for heritage digital twin. *Heritage Science*, *12*(1), 1-17.
18. Tan, G., Zhu, J., & Chen, Z. (2024). Deep learning based identification and interpretability research of traditional village heritage value elements: A case study in Hubei Province. *Heritage Science*, *12*(1), 1-17.

19. Sasithradevi, A., Chanthini, B., Subbulakshmi, T., & Prakash, P. (2024). MonuNet: a high performance deep learning network for Kolkata heritage image classification. *Heritage Science*, *12*(1), 1-14.
20. Net, F., Folia, M., Casals, P., Bagdanov, A. D., & Gómez, L. (2025). Eufcc-340k: A faceted hierarchical dataset for metadata annotation in glam collections. *Multimedia Tools and Applications*, *84*(32), 39259-39282.
21. The Metropolitan Museum of Art. (n.d.). *The Metropolitan Museum of Art Open Access* [Data set]. Kaggle. <https://www.kaggle.com/datasets/metmuseum/the-metropolitan-museum-of-art-open-access>