WORKING PAPER
N. 251
JUNE 2025

Text Analysis Methods for Historical Letters, The case of Michelangelo Buonarroti

Fabio Gatti

Joel Huesler

Università Bocconi
BAFFI
Centre on Economics, Finance and Regulation

Via Röntgen 1 | 20136 Milano – Italia | Tel +39 02 5836.5908/5564
baffi@unibocconi.it | www.baffi.unibocconi.eu

Text Analysis Methods for Historical Letters, The case of Michelangelo Buonarroti*

Fabio Gatti† Joel Huesler‡

June, 2025

Abstract

The correspondence of historical personalities serves as a rich source of psychological, social, and economic information. Letters were indeed used as means of communication within the family circles but also a primary method for exchanging information with colleagues, subordinates, and employers. A quantitative analysis of such material enables scholars to reconstruct both the internal psychology and the relational networks of historical figures, ultimately providing deeper insights into the socio-economic systems in which they were embedded. In this study, we analyze the outgoing correspondence of Michelangelo Buonarroti, a prominent Renaissance artist, using a collection of 523 letters as the basis for a structured text analysis. Our methodological approach compares three distinct Natural Language Processing Methods: an *Augmented Dictionary* Approach, which relies on static lexicon analysis and Latent Dirichlet Allocation (LDA) for topic modeling, a *Supervised Machine Learning* Approach that utilizes BERT-generated letter embeddings combined with a Random Forest classifier trained by the authors, and an *Unsupervised Machine Learning* Method. The comparison of these three methods, benchmarked to biographic knowledge, allows us to construct a robust understanding of Michelangelo's emotional association to monetary, thematic, and social factors. Furthermore, it highlights how the *Supervised Machine Learning* method, by incorporating the authors' domain knowledge and understanding of documents and background, can provide, in the context of Renaissance multi-themed letters, a more nuanced interpretation of contextual meanings, enabling the detection of subtle (positive or negative) sentimental variations due to a variety of factors that other methods can overlook.

*Preliminary draft, please do not quote without permission of the authors. We would like to thank Eric Strobl for his valuable comments during the process

†University of Bern, Switzerland & Baffi Center, Bocconi University, Italy; fabio.gatti@unibe.ch

‡University of Bern, Switzerland; joelhuesler@gmail.com

1 Introduction

Accurately reconstructing the emotional and thematic landscapes of the past from primary sources is a significant yet challenging undertaking in historical research. Letters, in particular, offer unparalleled insights by uniquely combining personal sentiment, self-presentation and transactional details that are rarely preserved elsewhere. However, their inherent complexity, characterised by antiquated vocabulary, evolving genres and highly specific socio-economic contexts , complicates large-scale, systematic interpretation. This article investigates whether contemporary Natural Language Processing (NLP) pipelines, specifically *Supervised* methods based on Large Language Models like BERT, can overcome these challenges to retrieve historical nuances more faithfully than traditional dictionary approaches.

We address this by applying three competing NLP workflows on a corpus of 523 letters written by Michelangelo Buonarroti (1475–1564), one of the most important artists in Renaissance Italy. Our core finding is that BERT's generated embeddings combined with Random-Forest architecture significantly outperforms, in the context of Renaissance multi-themed letters, both a classic *Augmented Dictionary* (based on static lexicon and Latent Dirichlet Allocation) baseline and an "off-the-shelf" *Unsupervised* BERT-based pipeline in capturing historically plausible sentiment and topic variation. These results contribute directly to ongoing debates regarding the promises and pitfalls of machine learning in the humanities, demonstrating how domain expertise can be productively integrated with state-of-the-art language models (Dobson, 2023; Izzidien et al., 2023; Beck and Köllner, 2023; Chun and Elkins, 2023; Ehrett et al., 2024; Tóth and Abdelzaher, 2024).

Early digital humanities research often relied on transparent, training-data-free methods such as word-list sentiment indices and n-gram frequencies (Silge and Robinson, 2017; Wehrheim, 2019). However, recent evaluations reveal that lexicons frequently misclassified metaphor, negation, and historical orthography, leading to attenuated or spurious findings in diachronic and multilingual corpora (Mello et al., 2023; Dennerlein et al., 2023). Transformer architectures offer a potential remedy, with contextual embeddings already improving lexicographic sense delineation (Tóth and Abdelzaher, 2024), cross-lingual sentiment detection (Chen et al., 2025), and multi-genre clustering (Gittel and Haider, 2025). Despite these advances, two critical issues persist for historical texts: the scarcity of labeled pre-modern training data and the "black box" nature of end-to-end transformer systems that can obscure historical interpretation (Dobson, 2023). Our *Supervised* pipeline directly addresses both problems by generating dense BERT embeddings, feeding them into a Random Forest classifier trained on a comparatively small, author-annotated benchmark dataset, and exposing variable importance scores for historical inspection. This approach aligns with recent successes in analyzing historical financial sentiment (Wehrheim et al., 2025) and interdisciplinary environmental corpora (Cheng et al., 2024; Suits and Moyer, 2024), while mitigating the opacity criticized by

Dobson (2023). Our main contribution lies in the comparison and evaluation of state-of-the-art Natural Language Processing tools applied to the specific case of historical letters, a type of document which lack a clear, singular thematic focus and instead interweave work-related, financial, familial, and social elements. These letters indeed present a dual layer of complexity. First, their content is inherently multi-thematic, resisting straightforward classification. Second, their vocabulary can departs from contemporary linguistic norms. Completing these two challenges, there is the functional role that early-modern letters often played: beyond their personal and expressive value, they frequently served as financial instruments, allowing the transfer of funds between employers and employees or across bank accounts. This dual communicative and transactional role adds further historical and economic depth, underscoring the need for understanding how different text-analysis tool can be sensitive to both linguistic, contextual and financial elements.

The specific historical context explored in this paper is Renaissance Italy, an era characterized by unparalleled economic prosperity, political fragmentation, and a flourishing artistic culture. This epoch was shaped by powerful patrons such as the Medici family and the Church authorities, whose wealth and influence were used for artistic innovation (Goldthwaite, 1987; Burke, 1999; Hatfield, 2002). Michelangelo Buonarroti (1475–1564), widely celebrated as one of the greatest Renaissance artists, left behind an extensive corpus of correspondence that provides insights into his artistic projects, interactions, and emotional states (Vasari, 1568; Zoellner and Thoenes, 2019). His letters are an exceptional case study, offering detailed coverage of diverse topics, including artistic commissions, financial transactions, personal reflections, and social relations. They are particularly valuable because they capture not only the artist's personal mindset but also the broader socio-economic dynamics and patronage systems integral to Renaissance culture, providing quantifiable insights into a period of significant tension between artistic expression, economic necessity, and social aspirations (Hatfield, 2002; Zoellner and Thoenes, 2019). We draw on the 523 outgoing letters preserved in the Archivio Buonarroti and published in the edition of Barocchi (1965). Rich metadata, including recipient identity, location, and 320 explicit monetary amounts embedded in the text, allow us to embed the NLP outputs in an econometric framework. After standard preprocessing, we produce three parallel representations for every letter: (i) dictionary sentiment and LDA topic weights; (ii) BERT embeddings classified by Random Forests trained on 215 hand-labelled thematic phrases and 129 sentiment exemplars; and (iii) unsupervised BERT embeddings clustered via k-means and scored with DistilBERT sentiment. This tripartite design facilitates a rigorous evaluation of accuracy, robustness, and historical plausibility across methods, for the case of multi-themed historical letters, representing an evaluation strategy recommended by recent AI-assisted studies on local heritage books (Stelter and Biehler, 2025) and newspaper geocoding (Sun and Qin, 2025).

Our analysis yields three key findings. First, the *Supervised Machine Learning* method accu-

rately reproduces Michelangelo's qualitatively anticipated shifts in tone, from positive when writing to "*Lords*" (private, political and religious elite of the time) to cooler with clerical middlemen, a pattern previously noted but never quantified (Hatfield, 2002). Second, it uncovers a non-linear "money–mood" curve: small routine payments where associated to lower sentiment-tone, while very large (outgoing) sums (as buying raw marbles) with higher, directly reflecting Michelangelo's documented pride in monumental projects (Zoellner and Thoenes, 2019). Third, only the *Supervised* pipeline detects a statistically significant frustration signal in work-related passages, a nuance entirely missed by both the lexicon and the unsupervised models. These results underscore recent warnings that off-the-shelf embeddings can misinterpret historically contingent language (Chun and Elkins, 2023; Dodda and Alladi, 2024). In summary, our results enable a good evaluation of NLP tools for analysing Renaissance multi-themed letters, benchmarked against the interpretations of the artists' biographers (Condivi, 1553; Vasari, 1568; Zoellner and Thoenes, 2019) A *Supervised Machine Learning* pipeline that uses BERT embeddings with labelled Random Forests effectively captures subtle historical nuances, such as Michelangelo's specific emotional correlations to finances and patrons, which align with qualitative scholarship (Hatfield, 2002; Vasari, 1568). In contrast, a *Augmented Dictionary* baseline, based on static lexicon and LDA, blurs these distinctions, and a fully *Unsupervised* transformer-based approach which remains noisy and often insignificant. This pattern confirms that robust historical inference from deep representations necessitates domain-specific supervision. It reinforces recent calls in digital humanities to hybridize "black-box" models with expert knowledge for enhanced accuracy and interpretability (VanBinsbergen et al., 2024; Eang and Lee, 2024; Dobson, 2023). Our approach clarifies the fundamental trade-off: lexicons offer transparency but are fragile; unsupervised transformers provide broad recall but risk anachronism; our hybrid model delivers the "explainable middle ground" now advocated in fairness research evaluating bias via embedding distances (Izzidien et al., 2023).

The remainder of this paper is structured as follows. Section 2 provides a concise historical overview of Renaissance Italy and Michelangelo's life, contextualizing his correspondence. Section 3 details the dataset and descriptive statistics. Section 4 outlines our NLP methodologies. Section 5 presents econometric analyses comparing method performance. Section 6 discusses these results in depth, connecting them to broader historical understandings and methodological implications. Finally, Section 7 concludes by summarizing our methodological contributions and suggesting future research directions.

2 Historical Background

2.1 Economics of Art in Renaissance Italy

In sixteenth century the Italian Peninsula appeared as divided in Regional States, in the North the Duchy of Savoy, the Duchy of Milan, the Republic of Venice, the Republic of Genoa, the Republic and then Duchy of Tuscany, in the Center the State of the Church, and in the South the Neapolitan Kingdom (Burke, 1999; Malanima, 2018). Despite this political fragmentation, around year 1500, Italy was among the more densely populated and urbanized areas in the Continent (Malanima, 2018; Burke, 1999) and probably the wealthiest (Goldthwaite, 1987; Koch et al., 2024; Braduel, 1966). In the specific, the wealth of the Italian States came both from industrial production (mostly of textiles), and from trade and banking activity (Mauro, 1990; Goldthwaite, 1987, 2009; Caracausi and Jeggle, 2014), and indeed, in most of the urban centres of the Peninsula, this wealth concentrated in the hands of a urban elite made by merchants and bankers. This concentration of wealth led to an unprecedented increase in the demand for unnecessary goods such as luxury and art, according to a correlation between artistic activity and entrepreneurial outcome which has been showed in the nowadays contexts of San Francisco and New York (Borowiecki, 2021). A prime example of this is Florence, the hometown of Michelangelo, where the Medici family, enriched though banking, surged at being the de facto rulers of the town (Malanima, 1983; Morelli, 1976; DeRoover, 1948), while constituting a network of patronage which fostered the production of art by the most promising Italian artist of the time. The other place where Michelangelo spent most of his life and earned the majority of his wealth (Hatfield, 2002) was Rome, where the Popes effectively established an autocratic although efficient revenue-collecting state, amassing fiscal resources which was partly invested in promoting the arts (Goldthwaite, 1987; Vaughan, 1908). Renaissance Italy thus consisted of a geographically and politically fragmented landscape, where decades of economic expansion gave rise to urban, ecclesiastical, and private elites who gained political and institutional power and invested in an elaborate system of artistic patronage for purposes of self-promotion and legacy-building (Goldthwaite, 1987; Monfasani, 2015; Malanima, 1983). In this hierarchically segmented society, the private or ecclesiastical *Lords* represented the demand side of a growing art market, in which the rich and powerful sought to elevate their status and prestige through the commissioning of artworks (Burke, 1999; Burckhardt, 1860; Borowiecki et al., 2024).

On the supply side were the artists, typically categorized—following Vasari (1568)—as painters, sculptors, and architects (Chambers, 1970). These individuals generally worked as "craftsmen" (Barkan, 2010; Hatfield, 2002; Zoellner and Thoenes, 2019) under the *cottimo* system. This arrangement meant that artists, usually registered with a local guild following an apprenticeship (Burke, 1999; Etro, 2018), were paid per piece based on ad hoc contracts with patrons. They were

typically responsible for covering the costs of raw materials[1]. More importantly, this system enabled the development of an individual "brand" value: renowned artists could command prices far above their material costs (Etro, 2018). This explains why patrons were willing to pay a premium for works by masters like Raphael, Michelangelo, and Leonardo, whose reputations guaranteed quality. These masters cultivated a distinct sense of individuality, each painting "in his own style" (Burke, 1999, p.29), setting themselves apart from one another and from medieval traditions. Despite their fame, however, most artists never rose above the status of craftsmen and were rarely granted access to public office. Only a few—including Michelangelo, Filippo Brunelleschi, and Giotto—achieved political recognition in an otherwise rigidly hierarchical early modern society (Hatfield, 2002).

2.2 Life of Michelangelo Buonarroti

Michelangelo Buonarroti was born in the small town of *Caprese*[2], on 6 March 1475, where his father Ludovico, member of a family of the lower Florentine aristocracy, was temporarily appointed as magistrate (Zoellner and Thoenes, 2019; Symonds, 1911). After few months the family moved back to Florence and spent his childhood between the city and a small estate in Settignano, a few kilometers outside of Florence, where he grew up among local marble cutters and stonemasons, gaining familiarity with marble and its carving from a young age (Zoellner and Thoenes, 2019). After receiving a standard primary education, Michelangelo was employed at the art workshop of Domenico Ghirlandaio, initially as an errand boy in 1487 and then as an apprentice painter in 1488 (Zoellner and Thoenes, 2019; Symonds, 1911), to then be noted by Prince Lorenzo de' Medici who invited the young Michelangelo to train at his expenses in workshop organized by the *Magnifico* himself in the San Marco gardens. With the death of Lorenzo de' Medici in 1492, Michelangelo fled to Bologna, just before the Medici family's expulsion from Florence, which followed the Charles VIII of France's invasion of Italy and the establishment of a new Florentine government under Savonarola's rule. Returned to Flroence in 1495, he created a *Sleeping Cupid*, which was fraudulently sold in Rome as an antique to Cardinal Raffaello Riario for 200 ducats[3]. (Symonds, 1911). The powerful clergyman discovered the fraud but, impressed by Michelangelo's skill, decided to summon the sculptor to Rome and commissioned works for the Holy See (Symonds, 1911),marking

[1]Such as marble for sculpture, wooden structures for painting or carving, costly pigments like ultramarine and gold, and the wages of specialized collaborators such as carpenters and goldsmiths (Malley, 2005).

[2]*Caprese* is located roughly 60km south east of Florence.

[3]In this paper, we refer to the denomination of monetary values as originally recorded in historical sources, such as the letters of artists and their contracts (Bardeschi, 2005; Barocchi, 1965). Specifically, monetary values are typically expressed in ducats, florins, or scudi (shields), depending on the regional Italian state. For analytical purposes, we treat all these currencies as equivalent (1:1), since during the early modern period, the primary currencies of Rome, Venice, and Florence fluctuated within a relatively narrow range—generally between 0.5 and 2 in value relative to one another. This approximation is supported by available collections of historical exchange rates among these currencies (DeRosa, 1955; Denzel, 2010; DaSilva, 1969). Regarding their modern-day equivalents, we estimate that one unit of currency from this period (e.g., one florin or ducat) would correspond to approximately 1,000 U.S. dollars today, based on the typical monthly wage of a manual laborer at the time. On this basis, the total payments received by the artist over the course of his career—including salaries and piecework commissions—amount to around 50,000 gold ducats, which would be equivalent to approximately 50 million U.S. dollars in today's currency (Hatfield, 2002)

the start of a new phase in his career. During this period, Michelangelo opened a Roman bank account which allowed him to collect payments from his Roman patrons and to transfer funds to support his father and brothers in Florence. Indeed, although the Buonarroti family belonged to a (marginal) branch of the Tuscan aristocracy, they consistently lacked true financial stability, before the success of Michelangelo (Zoellner and Thoenes, 2019). For his masterpieces Michelangelo was indeed soon rewarded: 150 gold florins for a *Bacchus*, 450 ducats for the *Pietà*(Milanesi, 1875). The *Tondo Doni*, committed by a powerful Florentine merchant worthed him 140 ducats (Vasari, 1568) while for the *Madonna with Child*, sculpted for a Flemish family, he received 100 ducats (Vasari, 1568; Hatfield, 2002). For the *David*, sculpted between 1501 abnd 1504, Michelangelo received a monthly payment of six ducats plus a final payment of 400 ducats (Milanesi, 1875).

By the time he turned thirty, the Florentine was already at the forefront of the Italian artistic scene, to the point that he was "inundated with work" (Zoellner and Thoenes, 2019). Nevertheless, he had to personally bear the costs of raw marble for some projects (Ciulich, 2005; Zoellner and Thoenes, 2019), while others were completed with significant delays [4] or never completed at all [5]. Between 1508 and 1512, he created the most celebrated fresco of the Italian Renaissance: the ceiling of the *Sistine Chapel*, for which the arist recived in total 3'000 ducats but had to bear the expenses for pigments and the wages of his collaborators (Hatfield, 2002). In the 1520s Michelangelo was then involved in several ambitious projects blending architecture and sculpture, like the Tomb of Lorenzo de' Medici, in a period which coincided with a profound personal loss, the death of his father, his closest familial tie in Florence and the main anchor in his personal life[6] (Zoellner and Thoenes, 2019). In 1534, at the age of 60, Michelangelo moved permanently to Rome, driven by his affection nobleman Tommaso de' Cavalieri other than by the call of the Popes, with Paul III who eventually appointed the Florentine as "Supreme Architect, Sculptor, and Painter to the Vatican Palace," granting him a monthly salary of 100 *scudi*, effectively turning his identity as a craftsman into that of a prestigious, salaried official of the Vatican (Hatfield, 2002). As an employee of the Pope, Michelangelo created two major works of his later career: the *Giudizio Universale* in the Sistine Chapel (1536-1541) (Hatfield, 2002; Zoellner and Thoenes, 2019) and the project of the St. Peter's Basilica, which would only be completed after his death, while also working on the Farnese Palace and the Capitoline Hill (Zoellner and Thoenes, 2019). Absorbed by these giant projects, Michelangelo almost ceased executing smaller commissions for private clients[7], but more interestingly, started to create some sculptures for himself, like the *'Florentine' Pietà* and the *Pietà Rondanini* (Zoellner and Thoenes, 2019). Michelangelo died in Rome on February 18, 1564,

[4]like the Tomb of Julius II for sum of 10,000 ducats, or the*Biblioteca Laurenziana*, completed after his death(Hatfield, 2002; Parker, 2010)

[5](like the for 15 Figures for the Duomo of Siena, of which he just realized 4, or the Facade of the Church of Saint Lorenzo in Rome

[6]His mother, Francesca, had died when he was just six years old (R. F. Sterba, 1978).

[7]With the exceptions of a *brutus* for Niccolo Ridolfi (Zoellner and Thoenes, 2019).

at the age of 89, in a modestly furnished home, where his friends found only sketches on papers, unfinished sculptures, a few items of clothing and over 30 kilograms of gold coins, reportedly enough to purchase the Palazzo Pitti, Florence's most imposing residence (Zoellner and Thoenes, 2019). According to his biographers, Michelangelo's frugality sharply contrasted with the considerable wealth he accumulated through his artistic success (Condivi, 1553; Vasari, 1568). By 1509, he was already so wealthy that the Florentine government repeatedly compelled him to participate in bond issues to support the city's finances (Hatfield, 2002), yet he continued to live like a pauper (Zoellner and Thoenes, 2019), dressing shabbily and rarely bathing (R. F. Sterba, 1978).

3 Data and Methodology

3.1 Data

The primary dataset comprises approximately 1,500 letters (of which 523 sent by the artists, toward which we address our analysis, and 868 received by him) between 1496 and 1564. These letters, written in Renaissance Italian, are collected and transcribed by Foundation Memofonte[8] which collected them from multiple editions of the artist's correspondence, including Barocchi (1965), Barocchi (1967), Barocchi (1973), Barocchi (1979), Barocchi (1983), and Barocchi et al. (1988). On average, a letter consists of 215 words, ranging from as few as 14 words to a maximum length of 2,321 words, a disparity suggesting the existence of a diverse range of communicative purposes: while some letters are succinct and direct, likely concerning simple instructions or brief acknowledgments, others are extensive and elaborate, potentially involving complex financial transactions, detailed artistic instructions, or deeply personal reflections. In general, an inspection of the correspondence allowed us to understand how the majority of the letters are multi-themed, consisting in a mix of working, familial and social arguments, while also often allowing for the transfer of financials, according to the Early-Modern practice of the bills of exchange.

A preliminary keyword search in the *corpus* of letters reveals the presence of different semantically rich terms. For instance, the term *Pope* appears 213 times, while *cardinal/cardinals* and *Medici* are mentioned 29 and 23 times, respectively. Artistic references are also frequent, with *art/arts* occurring 25 times, *sculpture/sculptures* 9 times, and *Church/Churches* 19 times. The primary recipients of Michelangelo's letters are close family members, reinforcing the personal dimension of his correspondence. His nephew Leonardo is the most frequent addressee, receiving 209 letters, followed by his brother Buonarroto with 78 letters, and his father Lodovico with 46. Geographically, Michelangelo primarily wrote from Rome (386 letters), followed by Florence (83) and Bologna. Conversely, the letters were mostly directed towards Florence (395) and Rome (96), highlighting the centrality of these cities (not by chance the "capitals" of the Renaissance) in his

[8] https://www.memofonte.it/ricerche/michelangelo-buonarroti/carteggio-diretto

communication network. The richness of this collection offers a unique window into Michelangelo's communication networks, financial dealings, and emotional expressions over time. This information is systematically extracted and analyzed through a combination of text analysis and econometric techniques, as described in the following sections.

3.2 Original Source

As mentioned above, our dataset originates from Barocchi (1965, 1967, 1973, 1979, 1983) which should be considered a comprehensive collection of all existing letters sent and received by Michelangelo (Barocchi, 1965). The effort to compile the artist's correspondence began nearly a century earlier with the work of Milanesi (1875). The primary source of this correspondence is the *Archivio Buonarroti*, which holds the original letters both sent and received by Michelangelo. The *Archivio*, initially the private archive of the Buonarroti family, became public following the death of Michelangelo's last descendant, Cosimo Buonarroti, in 1858. It is worth noting that, after Michelangelo's appointment as "Supreme Architect of the Holy See" in the 1530s, the volume of correspondence preserved in the archive dramatically decreased. This decline likely occurred because much of the organizational and work-related tasks previously handled through correspondence were subsequently managed directly by Vatican clerks (Bardeschi, 2005).

3.3 Preparatory Steps: Language, Social Groups and Moneys

3.3.1 Translation to English

As first preparatory steps, we translated the letters into modern English using Google Translate. We chose to translate the documents in order to refer to a broader literature employing language models trained on the English language (Devlin et al., 2019; Sanh et al., 2019), as these models have also been trained on larger samples compared to those available for the Italian language, likely allowing for a more subtle contextual understanding. Indeed, we manually checked 5% of the letters, verifying that the translation was satisfactory and that emotionally important terms (both negative and positive) were correctly translated into equivalent terms, even when misspelled (like *ghodere* translated into *enjoy*, even though in correct modern Italian it should be written without the *-h*, i.e., *godere*). The majority of work-related terms (like *sculptori*, again misspelled from *scultori*, but correctly translated into *sculptors*) and people-related appellatives (like *reverendissimo* into *most reverend* or *Messer* into *Sir*) were also correctly translated from Renaissance Italian to modern English.

3.3.2 Classification of the correspondent

Michelangelo sent letters to 66 different individuals which we manually assigned to a relational-social class on the base of Milanesi (1875) Parker (2010) and Stone (1962). This classification allows us to investigate heterogeneity in the communication style and emotional tones expressed in his letters, as well as to quantify these differences econometrically. We identified five macro-categories to represent the main social groups: *Family*, *Artists & Friends*, *Lords*, *Clerks*, and *Workers*. The *Family* category includes Michelangelo's closest relatives, such as his father Ludovico, his brothers Buonarrotto, Giovan Simone, and Gismondo, as well as his nephew Leonardo. Within the *Artists & Friends* category, we collected notable figures considered peers and confidants of Michelangelo. This includes celebrated personalities such as the painter, architect, and writer Giorgio Vasari, the goldsmith and sculptor Benvenuto Cellini, the poet Pietro Aretino, and the close friend and presumed lover Tommaso Cavalieri. The *Lords* category encompasses the highest ranks of Renaissance society, including members of the Medici family, Popes, cardinals, and bishops. The *Clerks* group consists of clerks, bankers, and priests who acted as financial or organizational intermediaries between Michelangelo and the elite. Lastly, we classified low-rank clerks, marble suppliers, and Michelangelo's helpers under the category of *Workers*.

3.3.3 Classification of the Monetary Transactions

The letters in our dataset contain a total of 320 monetary values[9]. Each letter may include anywhere from 0 to 11 monetary references. To analyse the impact of these amounts on Michelangelo's emotional tone, we first classified each monetary value based on its relationship to the letter's author. Specifically, we distinguished between positive relationships (money that the writer receives or is expected to receive in the future, whether immediately or as credit) and negative relationships (payments made or owed by the writer, including debts). After carefully examining the letters and considering the structure of the banking market at the time—particularly the involvement of intermediaries and agents handling transactions on behalf of others (DeRoover, 1944, 1953; Mandich, 1953; Kohn, 2001)—we refined this initial binary classification (1 for positive, -1 for negative) into a more detailed framework. In this revised approach, we introduced an intermediate classification of 0.5 for cases where the writer has an indirect but positive relationship with the mentioned money (i.e. when the money beneficiary is positively associated with the letter's writer, the typical case is that of a relative receiving the sum). Conversely, we assigned -0.5 when someone connected to the writer incurs a financial obligation. Finally, we accounted for instances where a monetary value holds no direct or indirect significance to the author (i.e., when the letter describes a transaction between third parties) by assigning a value of 0.

[9]Including both the letters sent and those received by the artist

3.4 *Augmented Dictionary* Method

3.4.1 LDA-based Topic Modeling

As starting step to perform the automated-text analysis over our *corpus* of letters we chose a Topic Modelling approach which is able to discover textual themes without any preceding knowledge or manual annotation, proceeding in an unsupervised manner (M. L. Jockers, 2020; Silge and Robinson, 2017). We opted for Latent Dirichlet Allocation (LDA), the most popular method for fitting a topic model when "not knowing what we're looking for" in terms of themes and letters' classification. This method has been widely applied in numerous papers of Economic History, to extract meaningful insights from historical texts and narratives (Wehrheim, 2019; Ferguson-Cradler, 2023; Gillings and Hardie, 2023; Kerby et al., 2022; Roberts et al., 2024; Wehrheim, 2024).

As Topic Modeling is based on the assumption that the semantic meaning of a text is created by the joint distribution of words within it and that topics are generated according to a stochastic process based on two probabilities: the probability of observing each word within a topic and the probability of observing each topic within a document (a letter, in our case), Latent Dirichlet Allocation (LDA) is employed as a probabilistic language model. Starting from the probability distribution of words across documents and a predefined number of topics, LDA estimates the topic probability distribution over words and the document distribution over topics. The generative process assumes that the probability of observing a word depends on the topic distribution for a given letter and the word distribution for a given topic according to the joint distribution (Blei et al., 2003):

$$p(w, z|\alpha, \beta) = \prod_{d=1}^{D} \prod_{n=1}^{N_d} \sum_{k=1}^{K} p(z_{d,n}|\theta_d) p(w_{d,n}|\phi_k),$$

where $p(z_{d,n}|\theta_d)$ represents the topic distribution for document d, and $p(w_{d,n}|\phi_k)$ is the word distribution for topic k. The hyperparameters α and β control the sparsity of the topic and word distributions, respectively (Blei et al., 2003).

To implement the LDA model on our dataset and uncover the latent topics in Michelangelo's letters, we employed the *topicmodels* R package. Following best practices for text preprocessing, we systematically removed punctuation, numbers, stopwords, and capital letters from the text corpus. This step ensured that the text was standardized and free from noise, optimizing the performance of the topic modeling process. After text preparation and transformation into a Document-Term Matrix (DTM), we executed the LDA model multiple times with different starting parameters. Specifically, we experimented with varying the number of topics from 2 to 50. The optimal number of topics was then determined based on the interpretability and coherence of the topics generated, leading to the selection of a model with $k = 7$ topics.

Among the seven identified topics, two were excluded due to their lack of interpretability, as

they consisted solely of contextual and unordered words. Additionally, two closely related topics were merged to improve clarity and thematic coherence. This refinement process resulted in four well-defined topics, which we labeled as *Work*, *Family*, *Money*, and *Patrons*. Table 1 displays the five most frequent terms for each of these four topics.

Topic	**Top Five Terms**
Work	Work, Marbles, Florence, Send, Pope
Family	Lionardo, Florence, Buonarroto, Simone, Roma
Money	Hundred, Ducats, Said, Money, Gold
Patrons	Lordship, Messer, Magnificent, God, Lord

Table 1: Top Five most common terms for each Topic

We consider LDA to be appropriately categorized under our *Augmented Dictionary* Approach, as it aligns with the fundamental principles of a "Bag of Words" methodology. Specifically, LDA, like dictionary-based sentiment analysis[10], disregards contextual meaning and the sequential order of words, instead relying solely on the statistical distribution of individual words across documents. This characteristic justifies its inclusion in the broader framework of text analysis methods that operate independently of syntactic or contextual nuances, focusing purely on word frequency and co-occurrence patterns or on *lexicon*-assigned scores.

3.4.2 Lexicon-based Sentiment Analysis

After the unsupervised quantification of topics for each letter, the second step of our *Augmented Dictionary* Method consists in another "bag of word" procedure aimed at quantifying the sentiment of each letter, completing the procedure exemplified in Figure 1. In this respect, considering a text (letter in our case) as the combination of individual words, the sentiment content of the whole text is calculated as the sum of the sentiment content of the individual words (Silge and Robinson, 2017; M. L. Jockers, 2020). This method offers a quick and pragmatic means of capturing the emotional tone of a text, making it particularly appealing for large-scale historical text analysis (Kabiri et al., 2023). Its primary advantage lies in its simplicity, relying only on a pre-compiled dictionary that maps specific words to sentiment scores, which are then aggregated to produce an overall measure of emotional polarity (Silge and Robinson, 2017; M. L. Jockers, 2020).

In this paper we employ a lexicon-based method *augmented* with valence shifters as implemented by the *Senrimentr* package in *R* (Rinker, 2021), where the sentiment of each sentence, s, is computed as:

$$s = \sum_{i=1}^{n} \text{weight}(w_i) + \text{shift}(w_{i-1}),$$

where $\text{weight}(w_i)$ represents the predefined sentiment score of the word w_i as drawn from the

[10]More details on this is provided in the following sections.

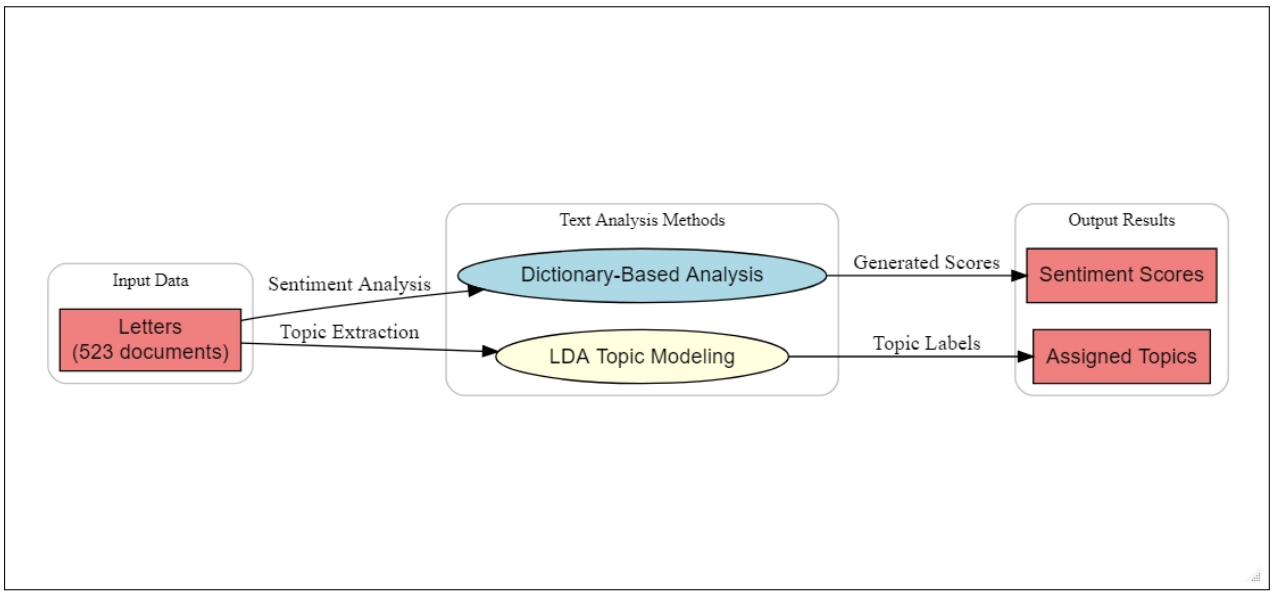

Figure 1: *Augmented Dictionary* Method

sentiment lexicon, and shift(w_{i-1}) captures the influence of *valence shifters*, such as *negators* (e.g., *not*), *amplifiers* (e.g., *very*), or *de-amplifiers* (e.g., *barely*), that modify the sentiment intensity of the following word.

One of the key strengths of this approach is its ability to interpret simple contextual cues that alter sentiment, particularly negation. For example, in the phrase *not happy*, the valence shifter *not* effectively reverses the positive sentiment of the word *happy*, allowing the method to accurately capture the negative tone of the expression. This represents a significant improvement over basic lexicon methods that simply count occurrences of positive and negative words without contextual adjustment. Anyway even with his *augmented* nature, this approach captures only immediate shifts in sentiment polarity due to negation or intensification, it does not comprehend deeper contextual meaning, a limitation inherent to the "Bags of Word" methodology.

3.5 *Machine Learning* Method

3.5.1 BERT's generated Word Embeddings

After constructing an initial dataset comprising sentiment scores and topic shares for each letter using the *Augmented Dictionary* approach, we proceeded to apply *Machine Learning* techniques to develop a comparable dataset. In the specific we experimented the use of Large Language Models to process the letters written by Michelangelo and generate words' embeddings, i.e. real valued vectors that encode the semantic representation of words (Surdenau and Valenzuela-Escarcega, 2024; Leavy et al., 2019), representing each as a point in a multi-dimensional space . As word embeddings are mathematical representation proved to be highly effective in capturing the contextual meaning of words, they've recently being used in different Economic History studies (VanBinsbergen et al.,

2024; Charlesworth et al., 2022). To generate the word's embedding for our *corpus* we employed a pre-trained language representation model named *BERT*[11] (Devlin et al., 2019), which is based on a multi-layer Bidirectional Transformer (Vaswani, 2017). The bidirectional approach allows BERT to achieve deeper contextual understanding than traditional left-to-right or right-to-left models, and it's based on the *cloze* task procedure first introduced by Taylor (1953). Such language model has been trained on a large *corpus* of modern English raw texts (not labeled by any human), such as Wikipedia, resulting in a linguistic and semantic "knowledge" not specific to any particular domain.

More precisely, our methodology consisted in applying the *BERT-base-uncased* model (Devlin et al., 2018)[12], exploiting its broad knowledge of English language to our dataset, without training a model from scratch, while directly using it "off-the-shelf" to process our dataset and generate words' embeddings for each letter. Such model contains and encoder with 12 layers (i.e. Transformer blocks) and 768 hidden units per layer and can take an input of no-more than 512 tokens. For this reason our methodology's first step consisted at first in splitting the letters into chunks below 500 words(Sun et al., 2020). After such splitting we fed the chunks to the model to generate word-level embeddings, and for each word we kept and concatenated the vectors output of the last two layers (11^{th} and 12^{th}) in a unique "final vector" in order to enrich the representation and the detail for each word, which in the end happened to be encoded in a 1536-dimensional vector. After the generation of word-level embeddings we re-aggregated such vectors first at "chunk" and then at letters level (our "unit of analysis"). To aggregate the tokens-embeddings into letters embeddings meant to us averaging the single-words vectors to generate an "averaged" letter-level embeddings. We assumed so that while each word's encoded vector captures a (fractional) part of the letter's content, the average of these can represent an holistic representation of a letter's overall semantic content. By averaging, we also diminish the impact of any one token or outlier, this is useful for noisy historical text where a single archaic or misspelled word might otherwise unduly influence a sequence representation. After such averaging, each letter's aggregated embedding can be understood as a point in a high-dimensional semantic space, summarizing the text's content. For example, a letter discussing a Chapel will have an embedding located near other vectors about Churches and Religion, even if specific words inside it would differ.

We chose to apply the *BERT-base-uncased* on R using the *text* package (Kjell et al., 2023), and not to train an ad-hoc model nor to fine tune it for three main reasons: first, as our letters are translated into "modern" english, their vocabulary is not inertly different from the cross-corpus on which the baseline BERT was trained. The second reason is that we have very few elements with which to train or fine-tune a new model (523 letters). Lastly, we don't use BERT as our

[11] *BERT* stands for Bidirectional Encoder Representations from Transformers.
[12] https://huggingface.co/google-bert/bert-base-uncased

end-to-end tool for letter classification, but rather as a tool to generate input for our own random forest classifier. We rely on downstream processing to adapt BERT's embeddings to our specific analytical tasks and texts, thus completing our machine learning approach.

3.5.2 *Supervised Method*: Random Forests for Classification

Following the generation of embeddings through *BERT-base-uncased*, we employed a *Supervised Machine Learning* approach to classify the Michelangelo's letters according to both their thematic and sentiment content, completing our *Machine Learning* approach. The Language Model, as described in the previous section, transforms each letter into a dense 1536-dimensional vector, resulting in a dataset of size 523×1536. These embeddings encapsulate the semantic properties of each letter and are the object for the subsequent classification process. To perform the classification task, i.e. to assign each of these vectors thematic and sentimental scores, we opted for a *Random Forest* classifier (Breiman, 2001a), a robust ensemble learning algorithm widely recognized for its effectiveness in handling high-dimensional data and its capacity to mitigate overfitting. Recently, this supervised approach has been increasingly applied in Economics (Coulombe, 2024; Liu et al., 2025), demonstrating its versatility and predictive strength in complex datasets.

Random Forests operate by constructing a multitude of decision trees during the training phase, with the final classification outcome determined by aggregating the predictions of individual trees through majority voting. This ensemble approach enhances both the stability and accuracy of predictions while maintaining interpretability through variable importance measures (Breiman, 2001a). Our classification task was structured into two distinct components: *topic classification* and *sentiment classification*. As our dataset consisted in 523 unlabelled letters (we didn't supplied any manual-classification for the letters in the original dataset), for the purposes of the *topic classification*, we constructed an external labelled datasetconsisting of 215 short phrases, each manually annotated to correspond with one of five thematic categories previously identified through the LDA approach: *Family*, *Money*, *Patrons*, *Work*, and *Neutral*[13] where each category contains exactly 43 phrases, ensuring balanced representation across the topics[14] . These theme-categories were selected to reflect the topics emerging from Michelangelo's correspondence as underlined from the LDA-based topic modelling, in order to produce a classification as comparable as possible with that computed by the *Augmented Dictionary* approach. The same *BERT-base-uncased* model described previously was then applied to these phrases to generate "labelled" phrase-wise embeddings, resulting in a feature matrix of size 215×1536. For the *sentiment classification*, we analogously constructed a separate labeled dataset comprising 129 short texts, categorized as *Positive*, *Negative*,

[13] A selection of *Neutral* labeled phrases was included to represent LDA topics that were excluded due to their lack of clear association with a coherent theme.

[14] We personally labelled each phrase, according to the most frequent label assigned to each phrase by 5 different PhD students at university of Bern

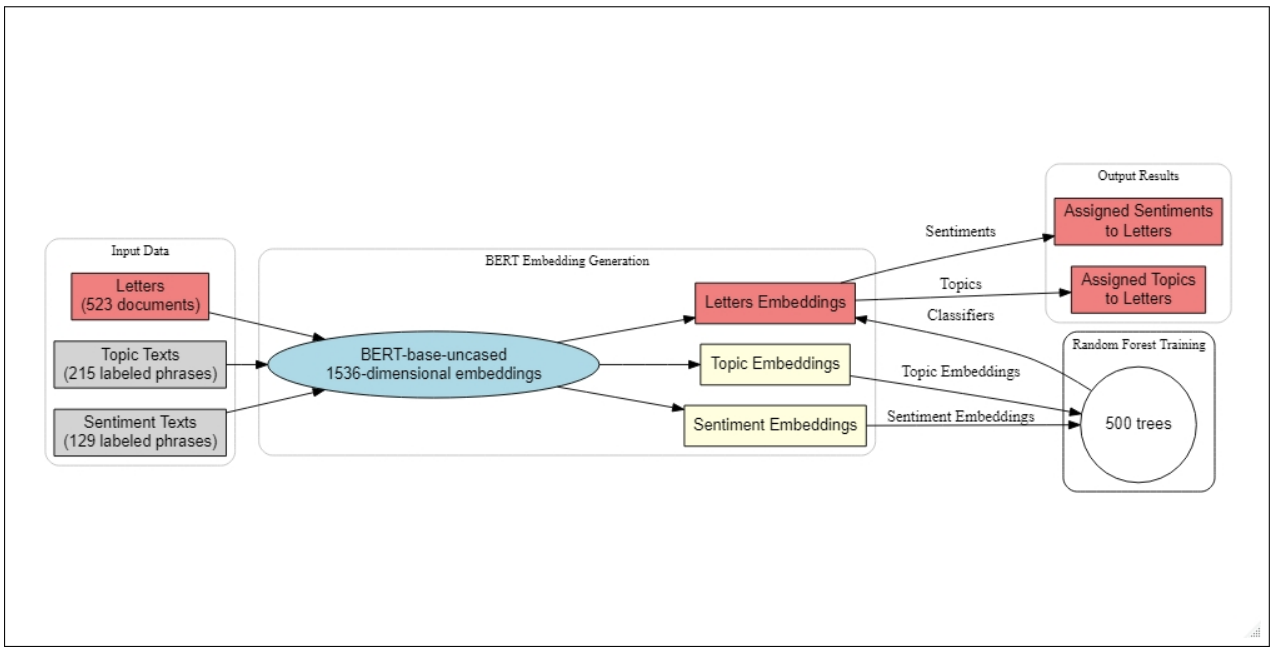

Figure 2: *Supervised Machine Learning* Method

or *Neutral*. As before, each class contains precisely 43 samples, guaranteeing balanced representation[15]. The embedding process mirrored that of the topic classification, producing a dataset of 129×1536 dimensions. Both these "labelled" datasets were then used to train two independent *Random Forest* classifiers each configured with 500 decision trees, trained one with the object of categorizing each embedding into a topic category and the other into one of three possible sentimental scores. We used the default hyperparameter settings of the *randomForest* package in R (Liaw and Wiener, 2002) which is based on Breiman (2001b), enabling the algorithm to optimise the number of nodes and split criteria dynamically during training. This approach ensures robustness against overfitting, making use of the model's inherent ability to handle high-dimensional feature spaces. More specifically, we trained the Random Forest on 95% of the respective 'labelled' datasets for both samples, using the remaining 5% as a test set. This resulted in Accuracies of 79% and 90.2% for the topic and sentiment classifiers, respectively, which are comparable to those already present in the literature (Leavy et al., 2019; Devlin et al., 2019; Eang and Lee, 2024; Reimers and Gurevych, 2023; Meija et al., 2023). At this point we employed the resulting two classifiers on our letters' *corpus*, resulting, for each letter, in two sets of probabilities, one for the topic classification and one for the sentiments. Figure 2 illustrates our entire *Supervised Machine Learning* approach, which we have structured to use supervised classification powered by BERT's ability to compute meaningful word embeddings.

[15]As for the topic-train set, the labels were given according to the majority of labels assigned by 5 PhD students from university of Bern

3.6 *Unsupervised Method*: K-Means clustering and *Distilbert*

After using the BERT generated embeddings to train a supervised classifier for topics and sentiments, we tried an *Unsupervised Machine Learning* method to further compare the results from different computer-based procedures. To generate topic categories by an unsupervised way, we applied *K-means clustering* (MacQueen, 1967; Murphy, 2012; Athey and Imbens, 2019) to the 523, 1536-dimensional, embeddings generated before. In order to allow for the presence of a "neutral" topic or of "not interpretable one", as done for the LDA and Random Forest Classifier, we clustered the embeddings in *K=5* groups[16]. Since the k-means algorithm assigns each observation exclusively to one cluster by looking at the minimum distance to each nearest centroid, this results in a hard classification in which each letter was exclusively assigned to a "thematic" cluster. However, given the multi-faceted and overlapping nature of Michelangelo's letters, we sought a more nuanced representation. We therefore computed the *Euclidean distance* from each letter to all five cluster centroids and transformed these into *topic probabilities* using an inverse-distance weighting scheme. in this way we assigned higher probabilities to closer clusters while still allowing for not-null probabilities for the others, providing a more flexible and expressive view of each letter's thematic positioning within the cluster space. At this point each letter is associated with a probability of belonging to each one of the five uncovered thematic clusters, configuring a set of data compatible to direct comparison with those estimated before through Latent Dirichlet Allocation and Random Forest Classification.

Also for the computation of the emotional tone of each letter, we re-quantifiyed it in an *Unsupervised Machine Learning* way, in the specific we employed the pre-trained LLM model *distilbert-base-uncased-finetuned-sst-2-english* (Sanh et al., 2019)[17], which was trained on a corpus of sentences extracted from movie reviews (Pang and Lee, 2005) and classified binarily as positive or negative by 3 human judges. As before, because of the low dimension of our *corpus*, we didn't fine-tuned the model, but applied it "off the shelf" and directly implemented on the letters to produce as output a sentiment score for each letter between -1 and +1, conceptually equivalent to the scores generated by the `sentimentr`' R package. Figure 3 is the graphical illustration of the *Unsupervised Machine Learning* method.

3.7 Descriptive Statistics

The original dataset, first processed as outlined in the *Preparatory Steps* section and further enriched through the *Augmented Dictionary*, *Supervised Machine Learning* and *Unsupervised Machine Learning* procedures, appears then as accompanied with rich metadata. In addition to information

[16] As from the previous methods we diagnosed the presence of 4 consistent topics, namely `Work, Money, Family, Patrons`

[17] https://huggingface.co/distilbert/distilbert-base-uncased-finetuned-sst-2-english

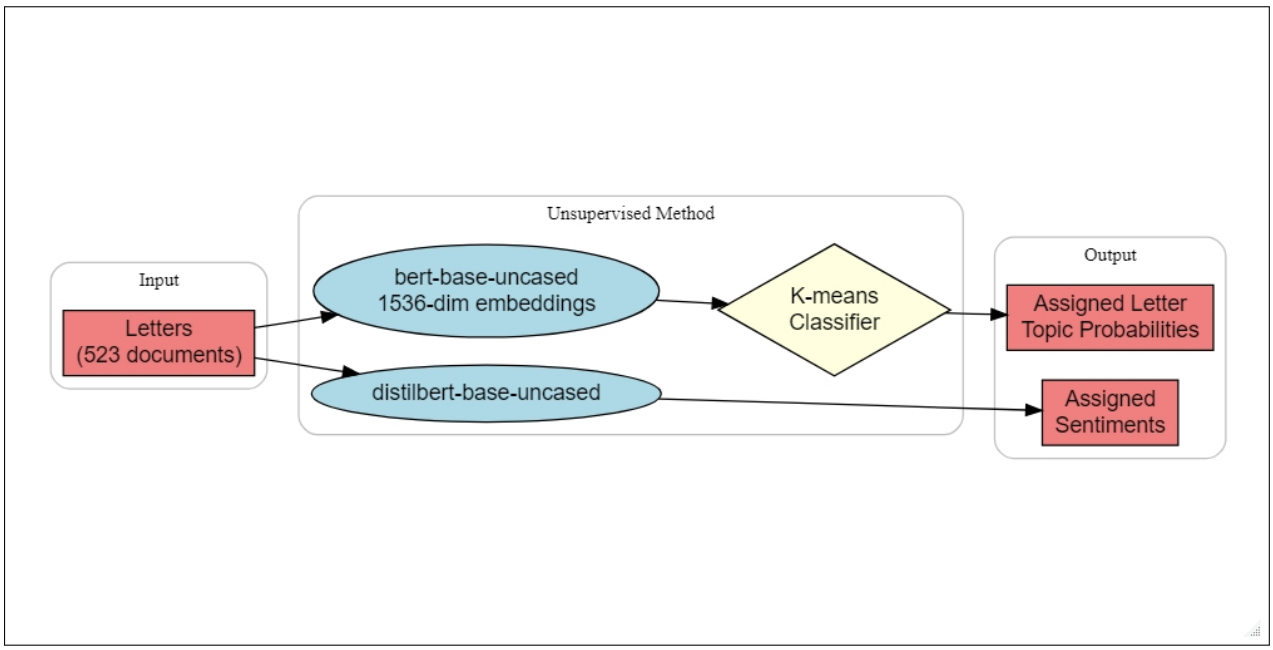

Figure 3: *Unsupervised Machine Learning* Method

on the time and location of each letter's dispatch, every entry includes the recipient's classification by *Social Group*. Furthermore, the dataset includes a detailed record of monetary transactions, complete with directional signs indicating the flow of financial exchanges. This metadata is then complemented by topic shares and sentiment scores derived from the *Augmented Dictionary* approach, alongside the topic probabilities and sentiment probabilities (both negative and positive) generated through the *Supervised Machine Learning* models, as well as with the probabilities of belonging to each one of the five thematic clusters, and a sentiment score assigned by the *Unsupervised Machine Learning* method. In this section we present graphical and descriptive evidences derived from the dataset, following the methodological steps outline above, in particular Figure 4 illustrates the geographical spread of Michelangelo's outgoing correspondence. The majority of the letters sent by the artist were dispatched in central and northern Italy, particularly along the route from Carrara, where he sourced marble, to Florence, and ultimately to Rome, where he spent most of his artistic career. The destination of such writings shows by the way a higher geographical reach, notably toward France, where he held a correspondence with King Francis I, which requested the Florentine artist to decorate his chapels.

In order to start a comparison between the topics identified in the three different *Methodologies* it's useful to start from the covariance-matrix depicted in Figure 5[18], which we used as guiding tool to interpret and assign semantic labels to the clusters identified through the *Unsupervised Machine Learning* method, by leveraging their relationship to those we computed via Latent Dirichlet Allocation and via the Random Forest Classifier. In Figure 5, topic probabilities derived by the

[18]The Figure shows the labels already "rescaled" by filtering out the "neutral topics" in the Random Forest classification and excluding the "discarded topics" in the LDA approach, and consequently also the less meaninful probabilities relative to the "extra cluster"cluster for the unsupervised method.

Augmented Dictionary Method are marked with the suffix _lda[19], those from the *Supervised Machine Learning* approach with _sup, and those from *Unsupervised Machine Learning* with _unsup. The matrix reveals a consistent cross-method alignment across all four topic categories: `Work`, `Money`, `Family`, and `Patrons`. Observing the values on and around the diagonal, we note a general agreement among the three methods in recognizing and clustering thematic content. However, the degree of alignment varies: the topic `Patrons` shows the highest average cross-method correlation (approximately 0.60), while `Work` presents the lowest agreement (average correlation around 0.34). In general, the consistently positive correlations across matching thematic areas—and the typically negative or near-zero correlations with non-matching ones—indicate that all three procedures tend to capture the intensity of thematic content in a similar direction. This provides preliminary evidence for the reliability and internal coherence of our triple-methodological approach, suggesting that the *Augmented Dictionary* as well as the two *Machine Learning* procedures converge in their representation of underlying semantic structures.

Moreover, Figure 6 presents the corresponding correlation analysis for sentiment detection. Unlike topic modeling—where outputs are represented as values between 0 and 1 interpreted as topic shares—the sentiment scores generated by our *Augmented Dictionary* approach rely on the `sentimentr` package in `R`. This method assigns a unique sentiment value to each letter, ranging from –1 (strongly negative) to +1 (strongly positive). For the purposes of this comparison, we normalized these values to a [0,1] scale and refer to this variable as A_D in the matrix. The *Supervised Machine Learning* method, analogous to topic classification, outputs two separate probabilities per letter: one for positive sentiment (ML_+) and one for negative sentiment (ML_-). From these, we derived an "overall" score, ($ML_=$), computed as the normalized difference between the two probabilities, providing a single [0,1] metric summarizing overall sentiment direction. Finally, the *Unsupervised Method* incorporates sentiment scores derived from the *distilbert-base-uncased-finetuned-sst-2-english* model. These were likewise rescaled to the [0,1] interval and denoted as (ML_{unsup}) in the matrix.

From Figure 6, we observe a general agreement among the three methods in how they quantify sentiment across the letters. In particular, there is a strong positive correlation between the overall supervised sentiment score ($ML_=$) and the probability of (supervised) positive sentiment (ML_+), (correlation around 0.93), confirming that letters assessed as more positive by the model tend to be associated with lower negative sentiment probabilities [20]. This is further supported by the strong negative correlation between $ML_=$ and ML_- (-0.93), indicating that the presence of negative sentiment cues directly reduces the overall sentiment score. A more moderate, yet still meaningful, alignment is observed between the dictionary-based sentiment score (A_D) and the unsupervised

[19] as for "Latent Dirichlet Allocation"

[20] this was indeed the underlying assumption in building the supervised overall sentiment score as the difference between the positive and the negative probabilities

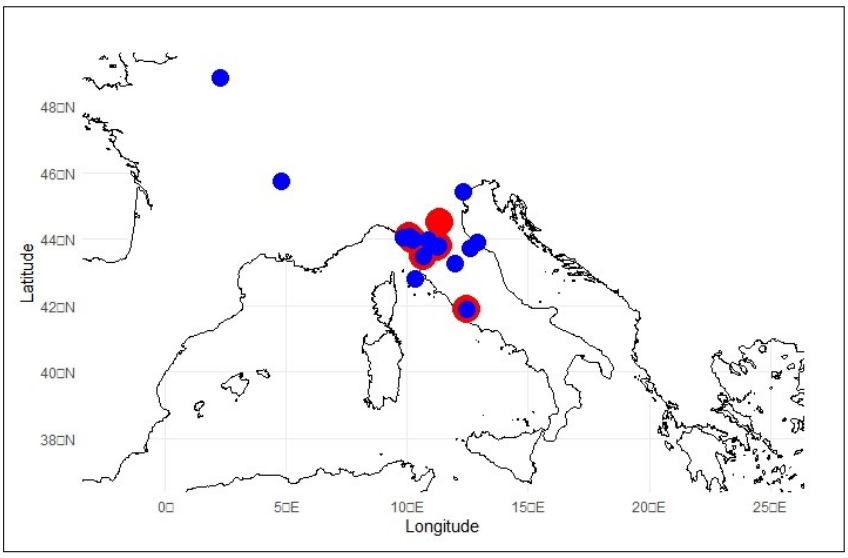

Figure 4: Geography of Michelangelo's Letters. In Red: origin, In Blue: destination

transformer-based score (ML_{unsup}), with a correlation of 0.30. Despite being grounded in fundamentally different methodologies–lexicon-based rule extraction versus Large Language Model-this correspondence suggests a convergent assessment of emotional tone across the letters.

In order to further the analysis of the correspondence as well as the comparison of the three text-analysis methods here employed, we intertwined *Sentiment Scores*, `Topic Shares` and *Social Class* for the thee Methodologies in Figures 7 (where sentiments are computed via *sentimentr* package in R, and the topics via Latent Dirichlet Allocation), Figure 8 (where sentiments displayed are the probabilities generated by the Random Forest for positive sentiment, ML_{+}, and the topic are as well generated by a *Supervised Machine Learning Approach*) and Figure 9 (where sentiments are generated by the *distilbert* LLM, and topics are unsupervisedly clustered) .

Several analogies emerge when comparing the three Figures. One clear similarity is that in all three methodologies, letters addressed to *Lords* display the highest average sentiment across all topics, indicating a consistently positive tone when Michelangelo communicated with members of the elite. Moreover for all methods the letters towards the *Medium-High Society* (i.e. the Clerks which intermediate between the lords and the artist) show the lowest emotional tone, suggesting a happiness in Michelangelo in relating to the highest sphere of the societies, and a certain frustration in dealing over preactical or bureucratic tasks with their representant over concrete works. In all three Methods the *Workers* occupy an intermediate position, while the *Augmented Dictiornay* and the *Supervised Machine Learning* tools diverge notably when it comes to sentiment expression towards *Family*. In the *Supervised Machine Learning* computed results, the letters directed to family members exhibit the lowest average positive sentiment, a surprising finding that is not mirrored in the LDA-based analysis, but it's partly confirmed in the *Unsupervised Results*.

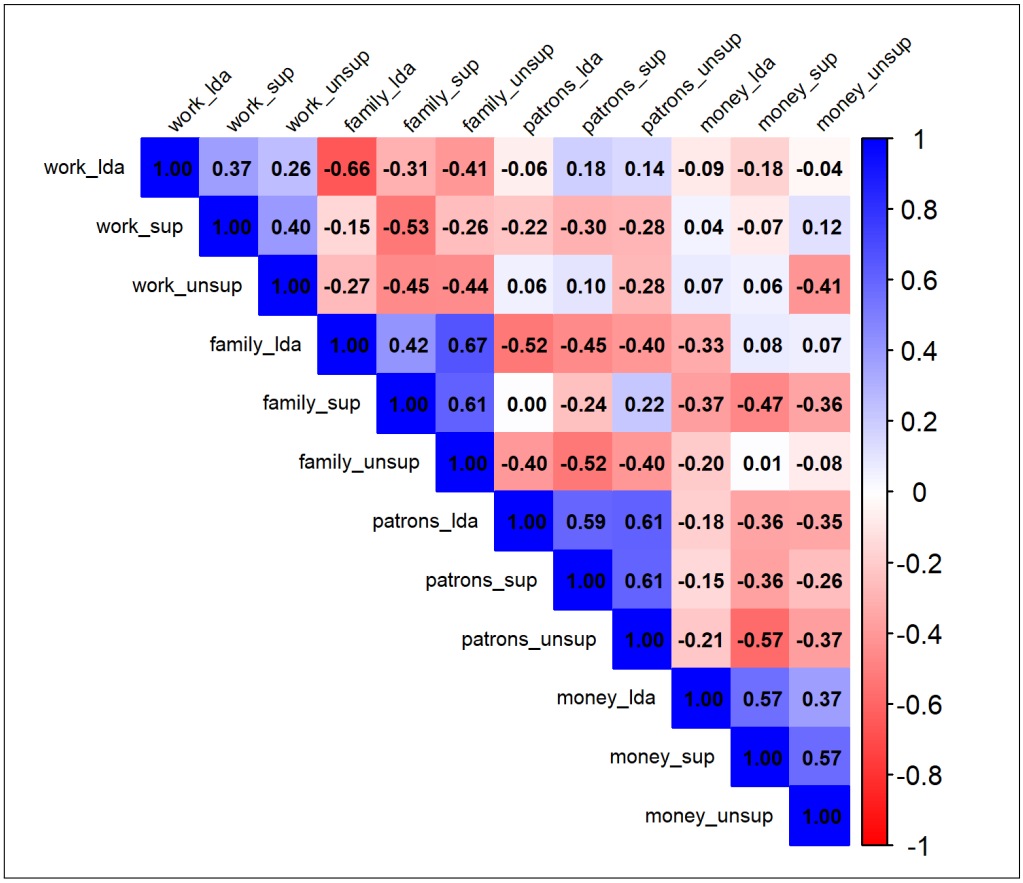

Figure 5: Correlations between *Augmented Dictionary*, *Supervised Machine Learning* and *Unsupervised Machine Learning* Topic Shares per letter.

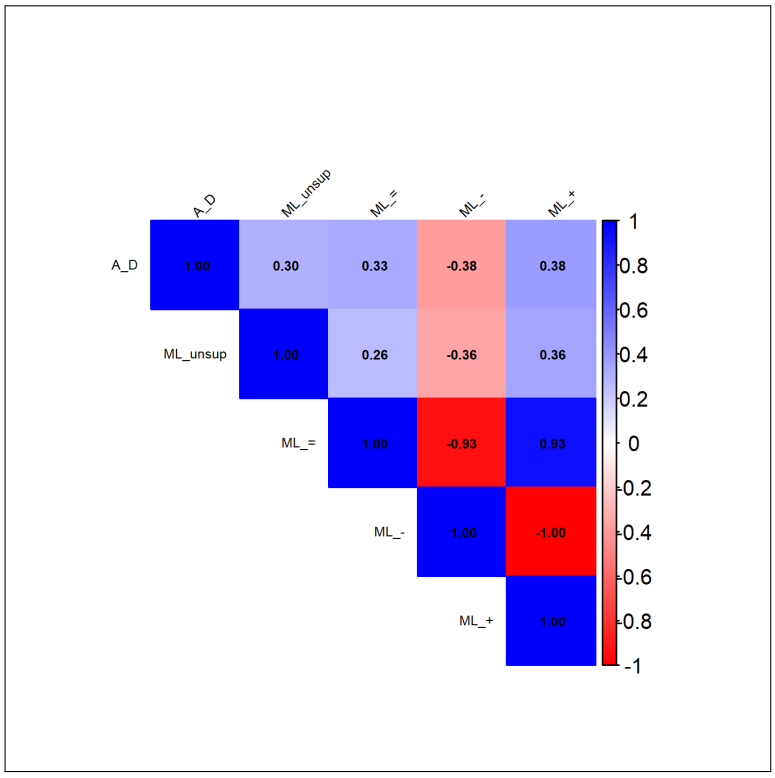

Figure 6: Correlations between *Augmented Dictionary*, *Supervised Machine Learning* and *Unsupervised Machine Learning* Sentiment Scores per letter.

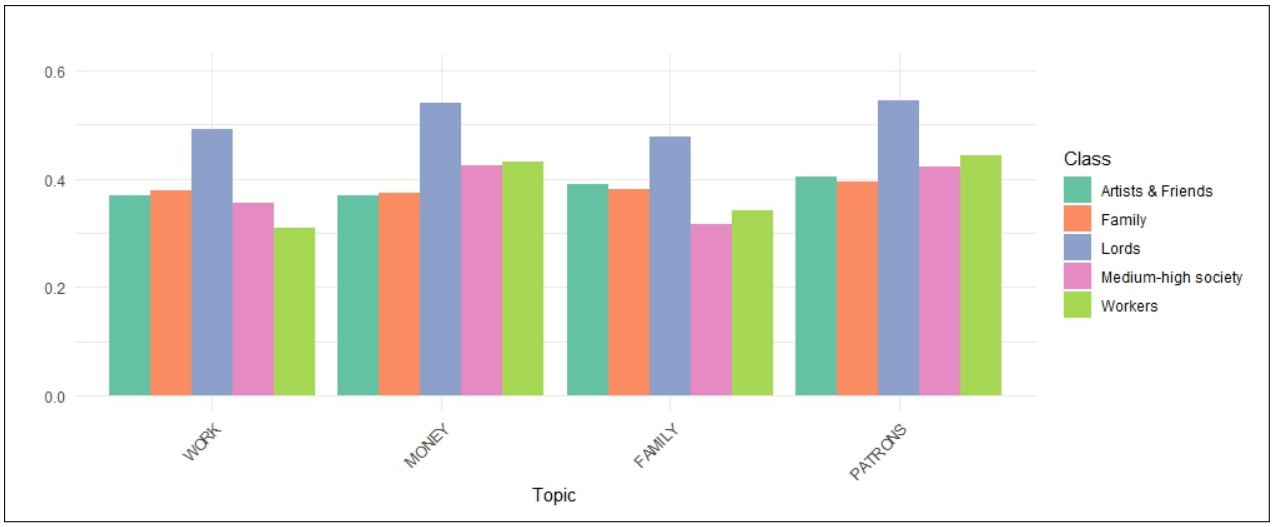

Figure 7: Average sentiment score (A_D) by topic and class (*Augmented Dictionary* Method)

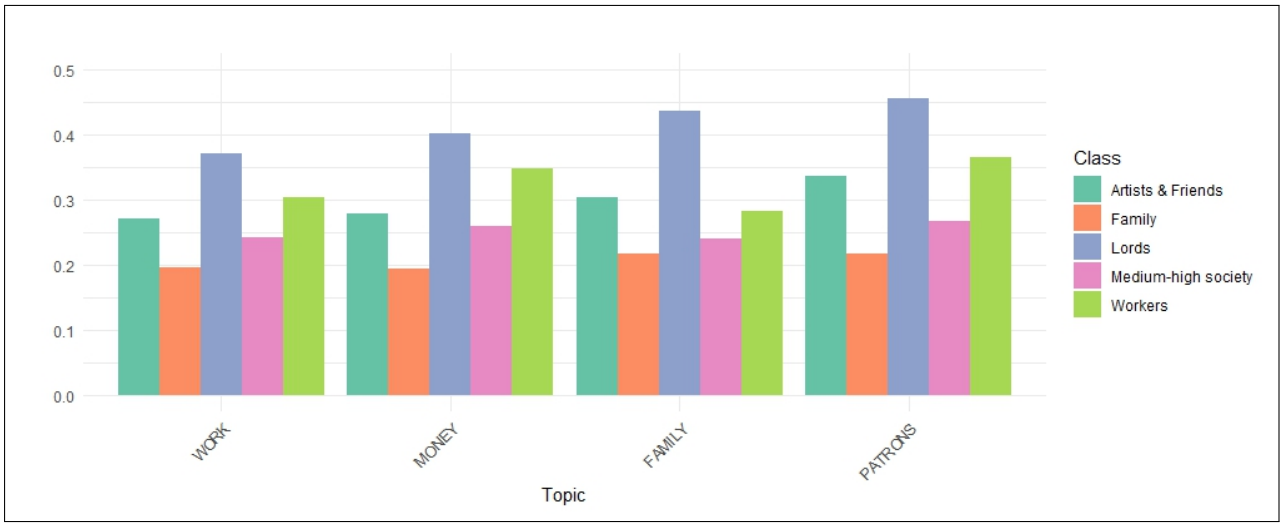

Figure 8: Average Positive sentiment ($ML\_(+)$) by topic and class (*Supervised Machine Learning* Method)

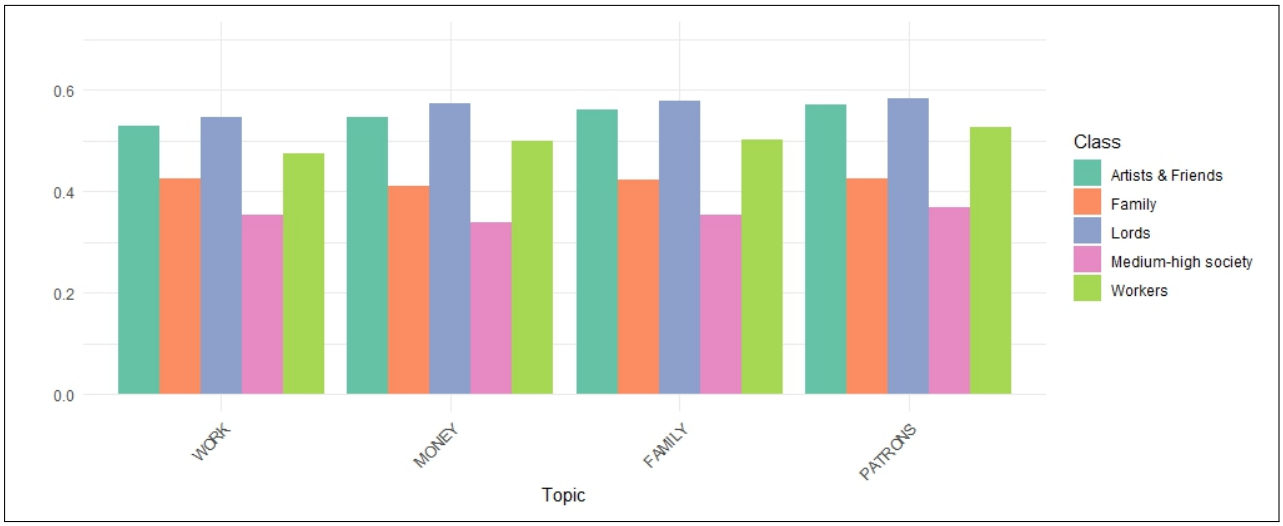

Figure 9: Average Unsupervised Sentiment score (ML_unsup) by topic and class (*Unsupervised Machine Learning* Method)

4 Results

4.1 Econometric Specification

We use Econometrics to test the analogies and differences of the three aforementioned *NLP* methods employed in this paper (the *Augmented Dictionary*, the *Supervised Machine Learning* and the *Unsupervised Machine Learning*) for the understanding of the interplay between sentiments, financial transactions and topics in the life of Michelangelo.

Equation 1 employs $\text{sentiment\_score}_{i,t}$ as the dependent variable, capturing the sentiment value assigned to letter i at time t. In the different specifications, this sentiment score is computed by *Augmented Dictionary*, the *Supervised Machine Learning* or the *Unsupervised Machine Learning* method. $\text{M\_sum}_{i,t}$ represents the sum of the absolute monetary values mentioned in a letter, irrespective of their sign. By contrast, $\text{M\_w+}_{i,t}$ denotes the weighted average sign (positive, negative, or neutral) of financial transactions, as perceived by the letters sender (Michelangelo). The interaction term $\text{M\_sum}_{i,t} \times \text{M\_w+}_{i,t}$ allows us to assess how both the magnitude and direction of financial flows jointly influence the sentiment expressed in a letter. $\text{Social\_Class}_{g,i,t}$ is a set of dummy variables identifying the recipient's social group: *Artists & Friends*, *Family*, *Lords*, *Clerks*, and *Workers*. All specifications include year fixed effects γ_{year} to control for time-specific unobserved heterogeneity.

$$
\begin{aligned}
\text{sentiment\_score}_{i,t} = \alpha + \; & \beta_1 \, \text{M\_sum}_{i,t} + \; \beta_2 \, \text{M\_w+}_{i,t} \\
& + \; \beta_3 \Big(\text{M\_sum}_{i,t} \times \text{M\_w+}_{i,t} \Big) \\
& + \; \sum_g \delta_g \, \text{Social\_Class}_{g,i,t} + \; \sum_g \theta_g \Big(\text{Social\_Class}_{g,i,t} \times \text{M\_sum}_{i,t} \Big) \\
& + \; \gamma_{\text{year}} + \varepsilon_{i,t}.
\end{aligned}
\tag{1}
$$

Equation 2 enriches the analysis by inserting Topics as independent variable in the specifications. Contrarily than in Equation 1 so, where the *Augmented Dictionary* and the two *Machine Learning* methods implementation regarded only the dependent variables, with the sentiments being quantified according to the three different approaches, here the differences arise also in the quantification of topics being estimated in some specifications via LDA, in other via the Random forest classifier and finally with unsupervised K-means clustering. In Equation 2 so, the $\text{sentiment\_score}_{i,t}$ is regressed against identifiers of the different $\text{Social\_Class}_{g,i,t}$ as well as against the different Topics (`Work`, `Money`, `Patrons`, `Family`) probabilities quantifiers $\text{Topic\_share}_{k,i,t}$, as well as against an interconnection of these with the sum, in absolute terms of the monetary values inside the letter i at time t.

$$\begin{aligned}\text{sentiment\_score}_{i,t} = \alpha + \sum_k \beta_k \, \text{Topic\_share}_{k,i,t} + \sum_g \delta_g \, \text{Social\_Class}_{g,i,t} \\ + \sum_k \beta_{k+4} \Big(\text{Topic\_share}_{k,i,t} \times \text{M\_sum}_{i,t} \Big) \\ + \gamma_{\text{year}} + \varepsilon_{i,t}.\end{aligned} \tag{2}$$

Together, these Equation models, in their different specifications, outlined below, a comprehensive framework for assessing the determinants of sentiment in Michelangelo's letters.

4.2 Regression Results

In the econometric results presented in Tables 2 and 3, the columns are labeled according to the type of $\text{sentiment\_score}_{i,t}$ used as the dependent variable. Specifically, *A_D* refers to the sentiment values computed using the *Augmented Dictionary* approach, as implemented by the *sentimentr* package in R. ($ML_{=}$) represents instead the specifications in which it's used as sentiment score the normalized overall score output of the *Supervised Machine Learning* method, computed as the difference between the likelihood that a letter conveys a positive emotional tone (ML_{+}) and the the probability estimation for a negative emotional tone in the letter (ML_{-}). Moreover, it is worth noting that the regression results using as dependent variable the two sentiment scores (ML_{+}) and (ML_{-}) are shown in Appendix for all the Equations and Specifications, in Table 6 to 9. Finally (ML_{unsup}) identifies the specification where the sentiment score are computed by the Large Language Model *distilbert*, as a part of our *Unsupervised Machine Learning* Method.

Moreover, it is worth noting that when including also topic shares (Tables 4 and 5), the *A_D*, $ML_{=}$ and ML_{unsup} specifications do not differ from each other only used dependent variable (i.e. the sentiment score per letter), but also in terms of the independent variables which represent the topics' shares for letters. Indeed in the (*A_D*) specifications, topic shares are derived using the Latent Dirichlet Allocation (LDA) method while in the ($ML_{=}$), the topics shares are generated as the probabilistic outputs of the ad-hoc Topic Random Forest Classifier. Lastly, in the ML_{unsup} model, topic shares correspond to the probabilities of belonging to the various cluster centroids obtained via an unsupervised *K-means* clustering algorithm applied to the BERT-derived embeddings.

In detail, Table 2 presents the estimation results of Equation 1, run in a simplified form. This specification includes only the monetary variables and year fixed effects as dependent variables, without controlling for the recipient's social class. The purpose of this "simplified" model is to capture a first heterogeneity in the correlations observed across different sentiment estimation methods. This allows us to have a first assessment in the consistency and divergence between the *Augmented Dictionary* and the *Supervised* and *Unsupervised Machine Learning*-based sentiment measures. In the *A_D* specifications, we observe that the weighted sign of the monetary transactions (M_{w_+}) within a letter is significantly and positively correlated with the artist's emotional expression, as

captured by the *sentimentr* package. Conversely, the absolute sum of the financial transactions (M_{sum}) is negatively correlated with sentiment, although this relationship is not statistically significant. The interaction term between the weighted sign and the absolute sum, as shown in the second column, is negatively and significantly correlated with the sentiment score. Comparing these results with the *Supervised Machine Learning*-generated sentiment scores, we observe notable similarities. In the ($ML_{=}$) specification, the sign of the transactions remains positively correlated with sentiment, while the absolute sum of the transactions is also negatively correlated. Unlike the A_D specification, however, the negative correlation between M_{sum} and sentiment is now statistically significant, suggesting that the *Supervised Machine Learning* -based sentiment analysis may better capture the emotional impact of large financial outflows. Further confirms of this can be seen in the Appendix tables 6 and which display the results for using (ML_{+}) and (ML_{-}), this latter showing, on principle, opposite signed coefficients respect to the former and to ($ML_{=}$) . In contrast, the ML_{unsup} specification—where sentiment is extracted using an unsupervised large language model—shows less definitive patterns. While the direction of two out of three coefficients aligns with the A_D specification, none of the results are statistically significant. This may indicate that "off-the-shelf" unsupervised sentiment models struggle to capture the emotional cues embedded in historical correspondence.

In Table 3, we present the estimation results for the fully specified version of Equation 1, which includes the social class of the recipient as an additional covariate. In certain specifications, we also introduce interaction terms between the recipient's social class and the monetary sums (M_{sum}). We begin by examining the baseline specifications (A_D), where the sentiment scores are derived using the augmented dictionary approach implemented in the *sentimentr* package. In the initial specification, which excludes the interaction terms between social class and monetary sums, the results mirror those of the simplified version. Specifically, the weighted sign of the monetary transactions (M_{w_+}) remains positively and significantly correlated with sentiment, while the absolute monetary sum (M_{sum}) remains insignificant. The interaction between the two variables, as before, is negatively and significantly correlated with the sentiment score, suggesting that the scale of transactions offsets the positive emotional influence of incoming funds when both are considered jointly. Turning to the social class variables, we observe that letters addressed to *Lords* and *Clerks* are associated with significantly higher sentiment scores. This result may indicate that Michelangelo expressed more positive sentiment in correspondence with individuals of higher status or administrative roles, potentially reflecting either respect or strategic politeness in his communication. When introducing the interaction terms between social class and monetary sums, interesting patterns emerge. Notably, the interaction terms for *Lords* $\times M_{sum}$ and *Family* $\times M_{sum}$ are both positive and statistically significant. This suggests that the emotional impact of financial transactions is amplified when the recipient belongs to these categories. The positive correlation for *Lords* may reflect Michelangelo's

dependence on or respect for his patrons, while the positive effect for *Family* could indicate a more intimate and emotionally charged nature of financial exchanges within familial correspondence.

Turning to the *Supervised Machine Learning*-based specifications, as with the results presented in Table 2, the overall sentiment specification ($ML\_(=)$) demonstrates multiple parallels with the augmented dictionary approach (A_D). Notably, the coefficients for the social classes *Lords* and *Clerks* remain positive and significant, indicating that Michelangelo's letters to these recipients are consistently associated with a higher sentiment score. This alignment suggests that the *Supervised Machine Learning* method successfully captures the positive emotional tone expressed towards individuals of high status and administrative influence, in line with the findings from the *Augmented Dictionary* method. However, an interesting divergence appears, however, in the case of letters addressed to *Family* members. The Machine Learning specifications, the sentiment tone in correspondence with family members is significantly higher, a pattern not captured in the A_D specifications. Again, for the specification relative to negative sentiments ($ML\_(-)$) and positive sentiments ($ML\_(+)$), the results are shown in the Appendix in Table 7, and mostly confirms the ($ML\_(=)$)'s specification's results. Confirming that the *Supervised Machine Learning* approach has a notable capacity in detecting the shift in the Michelangelo's emotional tone relative to outgoing financial transactions as well as the significant influence of social status on Michelangelo's emotional expression. Finally, the ML_{unsup} specifications again fail to produce statistically significant results, despite broadly confirming the direction of most coefficients observed in the previous models. Notable exceptions include the coefficient on M_{w_+}, which unexpectedly turns negative, suggesting an implausible aversion to incoming monetary flows, and some inconsistencies in the estimated effects for the *Family* social class. The only statistically significant coefficient in this specification concerns the *Worker* social class, which appears positively associated with sentiment. Interestingly, this effect was already present—but statistically insignificant—in both the *Augmented Dictionary* and *Supervised Machine Learning* specifications. These results further underscore the potential limitations of using pre-trained, off-the-shelf unsupervised models for capturing subtle emotional patterns in historical texts.

Table 4 presents the estimation results for a first specification of Equation 2, excluding interaction terms with social class and monetary variables. This regression table displays the correlation between sentiment scores, generated through different methods, and the thematic content of the letters, as captured by four main topics. In the baseline specifications (A_D), where sentiment scores are derived using the augmented dictionary approach implemented in the `sentimentr` package, the topic shares (`Work`, `Money`, `Patrons`) are computed through Latent Dirichlet Allocation (LDA). Notably, the *Patrons* topic emerges as the only positive and statistically significant predictor of sentiment. It is important to recall that four topics were initially identified, and the omitted category in this specification is `Family`, which serves as the baseline reference group.

In the *Supervised Machine Learning* specification sentiments and topics variables (`Work`, `Money`, `Patrons`) are derived from two ad-hoc Random Forest model that assigned probabilistic weights to each thematic category, reflecting the likelihood that a letter predominantly discusses one of these subjects. The results reveal a clear parallel between the overall ($ML\_(=)$) and the baseline model. Specifically, the *Patrons* topic, as identified by the Machine Learning approach, retains its positive and significant correlation with sentiment, mirroring the LDA-based results. Contrary to the A_D specifications, however, the Machine Learning-based models exhibit statistically significant coefficients for the *Work* and *Money* topics as well. In the Appendix Table 8 are shown the results for ($ML\_(+)$) and ($ML\_(-)$). In the *Unsupervised Machine Learning* specification, where sentiments are computed by the LLm model *distilbert*, and the topics are identified via *K-means clustering*, the results partially confirms those of the ($ML\_(=)$) specification, with a negative and significative coefficient relative to the `Topic`, and a positive, although not significant, one relative tol the `Lord`. By the way, contrary to the (A_D) and ($ML\_(=)$) models, the ML_{unsup} assign a negative and significant coefficient to the `Monetary` topic, producing a harldy conciliable difference respect the other two.

Table 5 presents the full estimation results for Equation 2, which includes social classes, monetary values, and thematic topics to comprehensively capture the determinants of sentiment in Michelangelo's letters. This specification aims to investigate how the artist's emotional tone is influenced by *to whom he is writing*, *what he is writing about*, and *whether money is involved*.

Focusing first on the A_D specifications, which rely on the augmented dictionary approach for sentiment analysis and Latent Dirichlet Allocation (LDA) for topic identification, we observe that the `Patrons` topic is significantly and positively associated with sentiment scores. In terms of social class, letters addressed to *Lords* are also associated with a higher sentiment score. Turning to the $ML\_(=)$ specifications, the results are broadly consistent with those of the A_D specifications. Specifically, the `Patrons` topic, identified here as a probabilistic output of the Random Forest model, continues to exhibit a positive and significant effect on sentiment. Similarly, the *Lords* class maintains a positive coefficient, although it does not reach statistical significance. Another analogy arises when examining the $ML\_(=)$ specifications for letters addressed to *Family*. Here, the sentiment tone is positive and statistically significant, indicating a warmer and more emotionally expressive tone in family correspondence. This pattern, although not significant, is confirmed in the baseline A_D model. Regarding the financial topic, the Machine Learning-identified `Money` topic shows a positive and significant correlation with sentiment in the $ML\_(=)$ specification, This finding is consistent with the A_D baseline, where the LDA-identified `Money` topic also carries a positive coefficient, albeit not statistically significant. As for the previous quations, the results relative to the $ML\_(+)$ and $ML\_(-)$, are in the Appendix in Table 9. The ML_{unsup} specifications continue to exhibit a general lack of statistically significant coefficients for both topics and interaction terms,

with the notable exception of a negative and significant coefficient for the `Money` topic—an effect not corroborated by the other two methodological approaches. Regarding the *Social Class* coefficients, the *Unsupervised Machine Learning* specification aligns with both the *Augmented Dictionary* and *Supervised Machine Learning* models in the direction of the estimated effects for each class. However, it also reveals a positive and statistically significant coefficient for the *Worker* class, a result that was previously present but not significant in the other two specifications.

Table 2: Regression results on Financials without Social Classes

Variable	A_D	A_D	$ML_{=}$	$ML_{=}$	ML_{unsup}	ML_{unsup}
M_{w_+}	0.0423**	0.0597***	0.0454**	0.0436**	-0.0262	-0.0126
	(0.0165)	(0.0183)	(0.0187)	(0.0211)	(0.0651)	(0.0736)
M_{sum}	-5.66E-07	-5.15E-07	-1.17E-06***	-1.17E-06***	-1.18E-06	-1.14E-06
	(3.55E-07)	(3.51E-07)	(4.01E-07)	(4.04E-07)	(1.40E-06)	(1.41E-06)
$M_{w_+} * M_{sum}$		-2.12E-05**		2.18E-06		-1.66E-05
		(1.02E-05)		(1.18E-05)		(4.11E-05)
Observations	523	523	523	523	523	523
R^2	0.292	0.2	0.402	0.403	0.213	0.214

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. In all the specifications, we controlled for Year Fixed Effects.

Table 3: Regression results on Financials including Social Classes

Variable	A_D	A_D	$ML_{=}$	$ML_{=}$	ML_{unsup}	ML_{unsup}
M_{w_+}	0.0469***	0.0164	0.0409*	0.0409**	-0.0149	-0.094
	(0.0174)	(0.0173)	(0.0217)	(0.0204)	(0.0747)	(0.0729)
M_{sum}	-4.92E-07	-4.79E-07	-9.34E-07**	-3.78E-07	-1.13E-06	-6.62E-07
	(3.34E-07)	(3.66E-07)	(4.16E-07)	(4.30E-07)	(1.43E-06)	(1.54E-06)
$M_{w_+} \times M_{sum}$	-2.17E-05**		2.76E-06		-2.52E-05	
	(9.46E-06)		(1.18E-05)		(4.06E-05)	
Family	0.00877	-0.0193	0.101**	0.113**	-0.00824	-0.129
	(0.038)	(0.0407)	(0.0473)	(0.0479)	(0.163)	(0.171)
Lords	0.240***	0.142*	0.103	0.232**	0.269	0.404
	(0.0565)	(0.0831)	(0.0704)	(0.0976)	(0.242)	(0.349)
Clerks	0.0985*	0.0349	0.0812	0.172*	0.285	0.131
	(0.056)	(0.0772)	(0.0698)	(0.0907)	(0.24)	(0.325)
Workers	0.0706	-0.0527	0.0622	-0.602	0.654**	0.125
	(0.0766)	(1.174)	(0.0955)	(1.379)	(0.329)	(4.934)
$M_{sum}\times$ Workers		0.000359		0.00627		0.00656
		(0.0113)		(0.0132)		(0.0474)
$M_{sum}\times$ Family		1.66E-05*		1.15E-06		4.36E-05
		(8.88E-06)		(1.04E-05)		(3.73E-05)
$M_{sum}\times$ Lords		7.18E-05**		-9.99E-06		2.76E-05
		(2.89E-05)		(3.40E-05)		(0.000122)
$M_{sum}\times$ Clerks		6.27E-05		-0.000128		3.47E-06
		(0.000102)		(0.00012)		(0.000429)
Observations	523	523	523	523	523	523
R^2	0.446	0.519	0.431	0.560	0.268	0.390

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. In all the specifications, we controlled for Year Fixed Effects.

Table 4: Sentiments- Topics, simple Correlations

Variable	A_D	$ML_{=}$	ML_{unsup}
`Work_topic`	0.0192	-0.365***	-2.648***
	(0.0274)	(0.0656)	(1.018)
`Patrons_topic`	0.117***	0.311***	1.4
	(0.0262)	(0.0716)	(1.021)
`Money_topic`	0.0255	0.625***	-5.254***
	(0.033)	(0.067)	(1.175)
Observations	523	523	523
R^2	0.189	0.388	0.238

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. All models include Year Fixed Effects. The ML_{unsup} column is based on unsupervised sentiment classification and does not yet include standard errors or R^2 values.

Table 5: Regression Results for Sentiments, Complete Models

Variable	A_D	A_D	$ML_{=}$	$ML_{=}$	ML_{unsup}	ML_{unsup}
`Work_topic`	0.105	0.0695	-0.143	-0.159	-1.977	-1.75
	(0.0945)	(0.0949)	(0.161)	(0.163)	(2.632)	(2.782)
`Patrons_topic`	0.251*	0.0991	0.631***	0.553***	-0.463	-0.156
	(0.144)	(0.153)	(0.187)	(0.202)	(3.04)	(3.165)
`Money_topic`	0.0679	0.0693	0.596***	0.654***	-3.724	-4.934*
	(0.0689)	(0.0724)	(0.104)	(0.117)	(2.364)	(2.6)
$M_{sum}\times$ `Work_topic`		-1.64E-05		-1.80E-05		-7.27E-05
		(1.19E-05)		(2.20E-05)		(0.00114)
$M_{sum}\times$ `Patrons_topic`		0.000215**		0.000109		-0.000502
		(8.76E-05)		(0.000119)		(0.00111)
$M_{sum}\times$ `Money_topic`		-5.23E-07		-4.76E-05		0.000566
		(1.71E-05)		(4.31E-05)		(0.000524)
Family	0.0795	0.0577	0.126***	0.122***	0.0307	0.0339
	(0.056)	(0.0556)	(0.0432)	(0.0436)	(0.198)	(0.1999)
Lords	0.248***	0.174***	0.0568	0.0286	0.287	0.248
	(0.0582)	(0.065)	(0.0624)	(0.0729)	(0.238)	(0.243)
Clerks	0.0815	0.0796	0.0303	0.0263	0.328	0.326
	(0.0574)	(0.0562)	(0.0586)	(0.0589)	(0.238)	(0.239)
Workers	0.0467	0.0516	0.117	0.119	0.575*	0.573*
	(0.0816)	(0.0801)	(0.0806)	(0.081)	(0.329)	(0.331)
M_{sum}	-4.92E-07	–	-2.78E-07		-1.15E-06	
	(3.57E-07)		(3.78E-07)		(-1.43E-06)	
Observations	523	523	523	523	523	523
R^2	0.42	0.456	0.606	0.611	0.284	0.293

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. All models include Year Fixed Effects. ML_{unsup} values are derived from an unsupervised classification and currently lack standard errors and R^2.

5 Discussion

Our empirical findings, presented in Tables 2 through 5, offer novel insights into Michelangelo's emotional life, which both align with and challenge existing historical literature on the Florentine master (Hatfield, 2002; Vasari, 1568; Zoellner and Thoenes, 2019; Condivi, 1553). Crucially, these results also provide a rigorous comparison of the sensitivity and interpretability of the three NLP tools employed. Regarding Michelangelo's emotional responses to monetary factors, both the *Augmented Dictionary* and *Supervised Machine Learning* methods indicate how a natural emotional uplift is associated to incoming payments, while a decline in sentiment is associated with outgoing ones. However, this intuitive pattern is significantly moderated by the absolute value of transactions. The negative coefficient on the transaction magnitude (M_sum) suggests that the emotional boost from receiving money weakens as the amount increases. Conversely, larger outflows are less emotionally negative than smaller ones. This nuanced relationship aligns with Michelangelo's biographers, who highlight his ambiguous relationship with money (Condivi, 1553; Vasari, 1568). For instance, sculptures, which Michelangelo considered his highest artistic expression (he famously signed his letters as "Michelangelo, sculptor"), typically fetched higher prices than paintings. The negative coefficient associated with large monetary sums likely reflects the artist's sorrow in detaching from these significant works, especially those into which he had invested considerable time. On the expenditure side, a similar mechanism may explain why smaller, routine outflows (e.g., for food or clothing), consistent with his documented frugality (Hatfield, 2002; Zoellner and Thoenes, 2019; Vasari, 1568; Condivi, 1553), generated more negative sentiment. In contrast, large payments, often for acquiring marble and preparatory materials for major artworks, may have symbolized artistic anticipation rather than financial loss, marking the transformative beginning of a creative process.

In contrast, the *Unsupervised Machine Learning* method proves notably less sensitive to emotional fluctuations caused by monetary values or their direction. While it consistently assigns a negative and significant coefficient to the `Money` topic in most specifications, this primarily reflects a different mode of quantifying emotions and themes in the correspondence, offering less granular insight into the artist's complex financial sentiments.

A second major contribution of our analysis concerns Michelangelo's interactions with individuals from different social backgrounds, particularly the political and religious elites of the time. All three methodologies consistently demonstrate (though the unsupervised method lacks statistical significance) a higher level of positive sentiment in letters addressed to individuals of high social standing (the *Lords*). This pattern reflects Michelangelo's aspirational relationship with the powerful elite of Renaissance Italy and aligns with the vertical segmentation of Renaissance society, where connections to influential figures like Top ranks of the Church and members of the Medici family represented economic opportunity, social recognition, and artistic validation (Goldthwaite,

1987; Hatfield, 2002; Etro, 2018). It also signals Michelangelo's ambition to transcend the conventional artisanal role ascribed to artists during the Renaissance (Vasari, 1568; Hatfield, 2002). Furthermore, our econometric models reveal a distinct emotional dimension in Michelangelo's familial correspondence. Letters addressed to family members convey positive sentiments across almost all specifications, particularly evident when discussing financial matters. This confirms that Michelangelo did not perceive financial support for his family as a burden, but rather as deeply rewarding, fulfilling his role as a dutiful son and brother committed to uplifting his family (Hatfield, 2002; Zoellner and Thoenes, 2019; Goldthwaite, 1987; Vasari, 1568).

In terms of thematic content, all NLP methodologies highlight a strong positive emotional association when Michelangelo discusses his `Patrons`. However, the *Unsupervised Machine Learning* method again lacks statistical significance in this finding, reinforcing its comparative insensitivity when applied to historical letters. This dominance of positive emotions towards his patrons, despite conflicts (often due to creative disagreements) noted in the literature (Parker, 2010; Zoellner and Thoenes, 2019; Hatfield, 2002; Vasari, 1568), is a key insight. Interestingly, the three NLP approaches produce notably divergent results regarding Michelangelo's emotional reactions to the theme of Work. While the lexicon-based approach identifies no statistically significant emotional response, both the *Supervised* and *Unsupervised Machine Learning* approaches convincingly demonstrate a negative emotional association with work-related topics, likely capturing subtle linguistic cues indicating stress, frustration, or fatigue. This divergence, when addressed through the lens of the artist's biographers (Condivi, 1553; Vasari, 1568; Zoellner and Thoenes, 2019), allow us to say how BERT's embedding based models show in this context an higher contextual understanding that may be overlooked by the lexicon-based ones (Silge and Robinson, 2017; Rinker, 2021; Wehrheim, 2019; M. L. Jockers, 2020; Gillings and Hardie, 2023; Ferguson-Cradler, 2023; Blei et al., 2003). Indeed, episodes of psychological and physical stressed did marked the professional life of Michelangelo, notably in the physical works while painting the Sistine Chapel ceiling and for the decades-long (and partially judicial) saga surrounding the realization of the Tomb of Julius II.

The divergent performance of the three pipelines is diagnostically revealing. Lexicon–LDA workflows continue to offer quick, transparent baselines and excel at flagging broad thematic clusters. Yet their static vocabularies and bag-of-words assumptions flatten the pragmatics of early-modern Italian prose, where irony, polysemy, and idiosyncratic spelling are rife (Murthy and Kumar, 2021). Unsupervised BERT embeddings, while context-sensitive, tend to drift into semantically diffuse clusters when deprived of labeled anchors. This produces sentiment scores highly sensitive to incidental lexical noise, a weakness reported in other historical applications (Zhang et al., 2022).

Supervised methods leveraging transformers-generated embeddings bridge these gaps by marrying a rich contextual understanding with historically informed labels. Even a modest hand-labeled set (215 thematic snippets; 129 sentiment snippets) trained a Random Forest classifier that gen-

eralized convincingly across five decades of letters. This result confirms experimental evidence from modern corpora showing that a few hundred annotations can unlock the full discriminative power of pre-trained Large Language models (Beck and Köllner, 2023). Our findings underscore that domain-specific supervision effectively operationalizes close-reading expertise at scale. Rather than replacing historians, ML pipelines extend their reach, provided they are trained on snippets historians have judged exemplary. Conversely, the disappointing performance of the unsupervised model cautions against deploying Large Language Models on archives without contextual tethering. Size alone does not substitute for situated knowledge.

6 Conclusion

This paper systematically evaluates competing NLP pipelines to identify the approach that most faithfully recovers the emotional and thematic texture of early-modern multi-themed correspondence. Evaluating Michelangelo's 523 letters, we compared (i) a transparent lexicon–LDA baseline, (ii) an unsupervised BERT combined with K-means workflow, and (iii) a hybrid approach that feeds BERT embeddings into a Random Forest trained on 344 hand-labelled snippets. Benchmarking the results against what stated by Michelangelo's biographers, the *supervised* method emerges as the clear winner. This hybrid NLP, allows to explain roughly 60% of the variance in letter-level sentiment—almost twice the fit of the lexicon model and three times that of the unsupervised transformer. These gains confirm a growing consensus that transformer representations reach their potential only when constrained by historically informed supervision (Beck and Köllner, 2023; VanBinsbergen et al., 2024).

Substantively, the supervised scores confirms and quantifies long-standing claims about Michelangelo's personality. He was financially cautious, emotionally invested in grand commissions, proud to support his relatives financially, deferential to patrons and periodically exasperated by routine workshop chores. The model also reveals an asymmetry that had previously gone unnoticed: small expenses were more painful than large ones. This opens up new avenues of research into the psychology of Renaissance artistic labour. For digital humanists, the lesson is practical. A modest annotation effort of a few hundred sentences can transform black-box embeddings into reliable historical measures, providing the rigour of distant reading while retaining the contextual subtlety of close reading. However, using pre-trained models 'as is' can lead to inaccurate inferences and undermine the credibility of quantitative cultural analysis. Future work could extend this hybrid template to other epistolary corpora and integrate fairness diagnostics (Izzidien et al., 2023), and explore multimodal linkages between text, material culture, and environmental data (Stelter and Biehler, 2025; Daheur and Le Noë, 2024). Supervised transformers are not a magic solution, but they are currently the most powerful tool we have for converting historical words into

historical evidence, based on domain knowledge.

7 References

Athey, S. and G. Imbens (2019): "Machine Learning Methods Economists Should Know About," *Annual Review of Economics*, 11.

Bardeschi, L. (2005): *I Contratti di Michelangelo*, Florence: Studio per edizioni scelte.

Barkan, L. (2010): *Michelangelo: A Life on Paper*, Princeton University Press.

Barocchi, P. (1965): "Il Carteggio di Michelangelo, Vol. I," .

——— (1967): "Il Carteggio di Michelangelo, Vol. II," .

——— (1973): "Il Carteggio di Michelangelo, Vol. III," .

——— (1979): "Il Carteggio di Michelangelo, Vol. IV," .

——— (1983): "Il carteggio di Michelangelo V," .

Barocchi, P., K. Bramanti, and R. Ristori (1988): "Il Carteggio Indiretto di Michelangelo," 2.

Beck, C. and M. Köllner (2023): "GHisBERT–Training BERT from scratch for lexical semantic investigations across historical German language stages," in *Proceedings of the 4th Workshop on Computational Approaches to Historical Language Change*, 33–45.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Borowiecki, K. J. (2021): "What makes an artist? The evolution and clustering of creative activity in the US since 1850," *Regional Science and Urban Economics*.

Borowiecki, K. J., C. M. . Gray, and J. Heilbrun (2024): *The Economics of Art and Culture*, Cambride University Press.

Braduel, F. (1966): *La Méditerranée et le Monde méditerranéen a l'époque de Philippe II*, New York: Harper Row.

Breiman, L. (2001a): "Random Forests," *Machine Learning*.

——— (2001b): "Random forests," *Machine learning*, 45, 5–32.

Burckhardt, J. (1860): *The Civilization of the Renaissance in Italy*.

Burke, P. (1999): *The Italian Renaissance, Culture & Society in Italy*, Princeton University Press.

Caracausi, A. and C. Jeggle (2014): *Commercial Networks and European Cities, 1400–1800*, London: Pickering Chatto.

Chambers, D. S. (1970): *Patrons and Artist in the Italian Renaissance*, MacMillan.

Charlesworth, T., A. Caliskan, and M. Banaji (2022): "Historical representations of social groups across 200 years of word embeddings from google books." *Proceedings of the National Academy of Sciences*.

Chen, R., Z. Zhong, X. Yuan, and H. Liu (2025): "Two sides of the same coin? Cross-linguistic sentiment comparison and thematic discovery of reader's reception of *Wolf Totem*," *Digital Scholarship in the Humanities*, 40, 40–53.

Cheng, K., S. Bensassi, R. J. R. Elliott, and E. Strobl (2024): "Constructing a county-level environmental events dataset for China during the Ming and Qing dynasties (1368–1911)," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 57, 123–145.

Chun, J. and K. Elkins (2023): "eXplainable AI with GPT4 for story analysis and generation: A novel framework for diachronic sentiment analysis," *International Journal of Digital Humanities*, 5, 507–532.

Ciulich, L. (2005): *I contratto di Michelangelo*, Studio per Edizioni Scelte.

Condivi, A. (1553): *Vita di Michelagnolo Buonarroti*.

Coulombe, P. G. (2024): "The macroeconomy as a random forest," *Journal of Applied Econometrics*.

Daheur, J. and J. Le Noë (2024): "Socio-ecological metabolism and rural livelihood conditions: Two case studies on forest litter uses in France and Poland (1875–1910)," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 57, 205–225.

DaSilva, J. (1969): *Banque et credit en Italie au XVIIe siecle*, Paris: Klincksieck.

Dennerlein, K., T. Schmidt, and C. Wolff (2023): "Computational emotion classification for genre corpora of German tragedies and comedies from 17th to early 19th century," *Digital Scholarship in the Humanities*, 38, 1466–1481.

Denzel, M. (2010): *Handbook of World Exchange Rates, 1590-1914*, Ashgate Publishing.

DeRoover, R. (1944): "What is Dry Exchange? A Contribution to the Study of English Mercantilism," *Journal of Political Economy*, 52.

——— (1948): *The Medici Bank: its organization, management, operations and decline*.

——— (1953): *L'´evolution de la lettre de change XIVe-XVIIIe si'ecles*, Paris: A. Colin.

DeRosa, L. (1955): *I cambi esteri del regno di Napoli dal 1591 al 1707*, Banco di Napoli.

Devlin, J., M. Chang, K. Lee, and K. Toutanova (2018): "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *CoRR*, abs/1810.04805.

——— (2019): "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.

Dobson, J. E. (2023): "On reading and interpreting black box deep neural networks," *International Journal of Digital Humanities*, 5, 431–449.

Dodda, R. and S. B. Alladi (2024): "BERT-based Document Clustering: Unveiling Semantic Patterns in 20News Group, Reuters, and BBC Sports Corpora," *Authorea Preprints*.

Eang, C. and S. Lee (2024): "Improving the Accuracy and Effectiveness of Text Clas-

sification Based on the Integration of the Bert Model and a Recurrent Neural Network (RNN$_{B}ert_{B}ased$),″ *applied sciences*.

Ehrett, C., L. Ghita, D. Ranwala, and A. Menezes (2024): "Shakespeare Machine: New AI-Based Technologies for Textual Analysis," *Digital Scholarship in the Humanities*, 39, 522–531.

Etro, F. (2018): "The Economics of Renaissance Art," *The Journal of Economic History, Vol. 78, No. 2 (JUNE 2018), pp. 500-538*.

Ferguson-Cradler, G. (2023): "Narrative and computational text analysis in business and economic history," *Scandinavian Economic History Review*, 71, 103–127.

Gillings, M. and A. Hardie (2023): "The interpretation of topic models for scholarly analysis: An evaluation and critique of current practice," *Digital Scholarship in the Humanities*, 38, 530–543.

Gittel, B. and T. Haider (2025): "The ongoing birth of the narrator: empirical evidence for the emergence of the author–narrator distinction in literary criticism," *Digital Scholarship in the Humanities*, 40, 509–528.

Goldthwaite, R. (2009): *The Economy of Renaissance Florence*, Baltimore: The Johns Hopkins University Press.

Goldthwaite, R. A. (1987): "The Economy of Renaissance Italy: The Preconditions for Luxury Consumption," *I Tatti Studies in the Italian Renaissance , 1987, Vol. 2 (1987), pp. 15-39*.

Hatfield, R. (2002): *The Wealth of Michelangelo*, Edizioni di Storia e Letteratura.

Izzidien, A., S. Fitz, P. Romero, B. S. Loe, and D. Stillwell (2023): "Developing a sentence level fairness metric using word embeddings," *International Journal of Digital Humanities*, 5, 95–130.

Kabiri, A., H. Landon-Lane, J. Tuckett, and R. Neyman (2023): "The role of sentiment in the us economy: 1920 to 1934," *The Economic History Review*.

Kerby, E., A. Moradi, and H. Odendaal (2022): "African time travellers: What can we learn from 500 years of written accounts?" *The Economic History Review*.

Kjell, O., S. Giorgi, and H. A. Schwartz (2023): "The text-package: An R-package for Analyzing and Visualizing Human Language Using Natural Language Processing and Deep Learning," *Psychological Methods*.

Koch, P., V. Stojkoski, and C. Hidalgo (2024): "Augmenting the availability of historical GDP per capita estimates through machine learning," *PNAS Research Article*.

Kohn, M. (2001): "Payments and the Development of Finance in Pre-Industrial Europe," *Dartmouth College Econ. Working Paper No. 01-15*.

Leavy, S., M. T. Keane, and E. Pine (2019): "Patterns in language: Text analysis of government reports on the Irish industrial school system with word embedding," *Digital Scholarship in the Humanities*.

Liaw, A. and M. Wiener (2002): "randomForest: Breiman and Cutler's Random Forests for

Classification and Regression," Tech. rep., r package version 4.7-1.1.

Liu, G., W. Long, and X. Luo (2025): "A Random Forest–Based Panel Data Approach for Program Evaluation," *Journal of Applied Econometrics*.

M. L. Jockers, R. T. (2020): *Text Analysis with R For Students in Literature*, Springer.

MacQueen, J. (1967): "Some Methods for Classification and Analysis of Multivariate Observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, CA: University of California Press, vol. 1, 281–297.

Malanima, P. (1983): *La Decadenza di un'Economia cittadina L'industria a Firenze nei secoli XVI-XVIII*, Bologna: Il Mulino.

——— (2018): "Italy in the Renaissance: a leading economy in the European context, 1350-1550," *The Economic History Review, Vol. 71, No. 1 (FEBRAURY 2018), pp. 3-30.*

Malley, M. O. (2005): *The Business of Art, Contracts and Commissioning Process in Renaissance Italy*, Yale University Press New Haven and London.

Mandich, G. (1953): *Mandich, Giulio. Le pacte de ricorsa et le marche italien des changes au 17e siecle*, Paris: Affaires et gens d´affaires.

Mauro, F. (1990): "Merchant communities, 1350-1750." *The Rise of Merchant Empires Long Distance Trade in the Early Modern World 1350–1750*, Cambridge University Press, 264.

Meija, J., M. Harvill, and M. Xue (2023): "Techniques for Extracting Meaningful BERT Sentence Embeddings for Downstream Tasks," *Stanford CS224N Default Project*.

Mello, C., G. S. Cheema, and G. Thakkar (2023): "Combining sentiment analysis classifiers to explore multilingual news articles covering London 2012 and Rio 2016 Olympics," *International Journal of Digital Humanities*, 5, 131–157.

Milanesi, G. (1875): *Le lettere di Michelangelo Buonarroti*, Firenze, Coi tipi dei successori Le Monnier.

Monfasani, J. (2015): "THE RISE AND FALL OF RENAISSANCE ITALY," *Aevum , Settembre-Dicembre 2015, Anno 89, Fasc. 3 (Settembre-Dicembre 2015), pp. 465-481.*

Morelli, R. (1976): *La seta fiorentina del cinquecento*, Milano: Giuffrè Editore.

Murphy, K. P. (2012): *Machine Learning: A Probabilistic Perspective*, Cambridge, MA: MIT Press.

Murthy, A. R. and K. A. Kumar (2021): "A review of different approaches for detecting emotion from text," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 1110, 012009.

Pang, B. and L. Lee (2005): "Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales," in *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL)*, Ann Arbor, Michigan: Association for Computational Linguistics, 115–124.

Parker, D. (2010): *Michelangelo and the Art of Letter Writing*, Cambridge University Press.

R. F. Sterba, E. S. (1978): "The Personality of Michelangelo Buonarroti: Some Reflections," *American Imago , SPRING-SUMMER, 1978, Vol. 35, No. 1/2,*.

Reimers, N. and I. Gurevych (2023): "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*.

Rinker, T. (2021): "Package 'sentimentr'," *Retrieved*, 8, 31.

Roberts, M., B. Stuard, and et. al (2024): "Structural topic models for open-ended survey responses," *American journal of political science*.

Sanh, V., L. Debut, J. Chaumond, and T. Wolf (2019): "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*.

Silge, J. and D. Robinson (2017): *Text Mining with R, A Tidy Approach*, O'Reilly.

Stelter, R. and R. Biehler (2025): "Data retrieval from local heritage books—Is artificial intelligence the solution?" *Historical Methods: A Journal of Quantitative and Interdisciplinary History*.

Stone, I. (1962): *I, Michelangelo, Sculptor*, Garden City, New York: Doubleday Company.

Suits, R. and E. Moyer (2024): "Estimating energy flows in the long run: Agriculture in the United States, 1800–2020," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 57, 242–251.

Sun, C., X. Qiu, Y. Xu, and X. Huang (2020): "How to Fine-Tune BERT for Text Classification?" *Chinese Computational Linguistics: 18th China National Conference*.

Sun, S. and X. Qin (2025): "Fine-grained extraction of geospatial and temporal information from Chinese historical newspapers," *Digital Scholarship in the Humanities*, 40, 601–616.

Surdenau, M. and M. Valenzuela-Escarcega (2024): *Deep Learning for Natural Language Processing*, Cambridge University Press.

Symonds, J. A. (1911): *The Life of Michelangelo based on studies in the archives of the Buonarroti family at Florence, Vol. 1*, University of Pensylvania Press.

Taylor, W. L. (1953): "Cloze Procedure: A New Tool For Measuring Readability," *Journalism Quarterly*.

Tóth, A. and E. Abdelzaher (2024): "BERT may help in lexicographic sense delineation," *International Journal of Digital Humanities*, 6, 291–311.

VanBinsbergen, J., S. Bryzgaova, M. Mukhopadhyay, and V. Sharma (2024): "(almost) 200 years of news-based economic sentiment," *National Bureau of Economic Research*.

Vasari, G. (1568): *Le vite de' più eccellenti pittori, scultori e architettori*, Lorenzo Torrentino.

Vaswani, A. (2017): "Attention Is All You Need," *Advances in Neural Information Processing Systems*.

Vaughan, H. M. (1908): *The Medici Pope*, London: Methuen.

Wehrheim, L. (2019): "Economic history goes digital: topic modeling the Journal of Economic

History," *Cliometrica*, 13, 83–125.

——— (2024): "Digital methods in economic history: The case of computational text analysis." *In Handbook of cliometrics*.

Wehrheim, L., J. Borst-Graetz, B. Liebl, M. Burghardt, and M. Spoerer (2025): "More than a feeling. Introducing an NLP-based media sentiment index for the Berlin Stock Exchange, 1872–1930," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, advance online publication.

Zhang, P., H. Zhao, F. Wang, Q. Zeng, and S. AMOS (2022): "Fusing LDA Topic Features for BERT-based Text Classification," .

Zoellner, F. and C. Thoenes (2019): *Michelangelo the complete Works*, Taschen.

Table 6: Regression results on Financials without Social Classes (ML_+ and ML_-)

Variable	ML_+	ML_+	ML_-	ML_-
M_{w_+}	0.0187	0.0191	-0.0444*	-0.0419
	(0.0139)	(0.0157)	(0.0238)	(0.0269)
M_{sum}	-1.98E-07	-1.98E-07	1.35E-06***	1.36E-06***
	(2.97E-07)	(3.00E-07)	(5.11E-07)	(5.14E-07)
$M_{w_+} * M_{sum}$		-4.07E-07		-3.10E-06
		(8.75E-06)		(1.50E-05)
Observations	523	523	523	523
R^2	0.381	0.381	0.387	0.387

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. In all the specifications, we controlled for Year Fixed Effects.

Table 7: Regression results on Financials including Social Classes (ML_+ and ML_-)

Variable	ML_+	ML_+	ML_-	ML_-
M_{w_+}	0.0133	0.0104	-0.0426	-0.0448*
	(0.0148)	(0.0142)	(0.0269)	(0.0257)
M_{sum}	-2.79E-07	-1.43E-07	9.92E-07*	3.79E-07
	(2.84E-07)	(2.99E-07)	(5.18E-07)	(5.42E-07)
$M_{w_+} \times M_{sum}$	-2.84E-06		-5.65E-06	
	(8.04E-06)		(1.46E-05)	
Family	-0.0384	-0.0174	-0.158***	-0.158**
	(0.0323)	(0.0333)	(0.0588)	(0.0604)
Lords	0.125**	0.200***	-0.0397	-0.15
	(0.048)	(0.0679)	(0.0875)	(0.123)
Clerks	0.0557	0.101	-0.0629	-0.146
	(0.0476)	(0.0631)	(0.0868)	(0.115)
Workers	0.164**	-0.999	0.042	0.0309
	(0.0652)	(0.96)	(0.119)	(1.74)
$M_{sum}\times$ Workers		0.0109		7.93E-05
		(0.00922)		(0.0167)
$M_{sum}\times$ Family		-1.35E-06		-2.47E-06
		(7.26E-06)		(1.32E-05)
$M_{sum}\times$ Lords		-2.31E-05		-4.36E-06
		(2.36E-05)		(4.29E-05)
$M_{sum}\times$ Clerks		-6.00E-05		0.00012
		(8.34E-05)		(0.000151)
Observations	523	523	523	523
R^2	0.501	0.599	0.444	0.557

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. In all the specifications, we controlled for Year Fixed Effects.

Table 8: Sentiments–Topics, simple correlations (ML_+ and ML_-)

Variable	ML_+	ML_-
`Work_topic`	-0.174***	0.339***
	(0.0587)	(0.0792)
`Patrons_topic`	0.650***	0.0836
	(0.064)	(0.0864)
`Money_topic`	-0.124**	-0.895***
	(0.0599)	(0.0809)
Observations	523	523
R^2	0.44	0.448

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. All models include Year Fixed Effects.

Table 9: Regression Results for Sentiments, Complete Models (ML_+ and ML_-)

Variable	ML_+	ML_+	ML_-	ML_-
`Work_topic`	-0.290**	-0.300**	-0.0314	-0.0183
	(0.125)	(0.127)	(0.204)	(0.206)
`Patrons_topic`	0.229	0.196	-0.641***	-0.565**
	(0.146)	(0.158)	(0.236)	(0.256)
`Money_topic`	-0.0528	-0.0219	-0.804***	-0.857***
	(0.0808)	(0.0911)	(0.131)	(0.148)
$M_{sum}\times$ `Work_topic`		-4.35E-06		1.99E-05
		(1.71E-05)		(2.78E-05)
$M_{sum}\times$ `Patrons_topic`		3.72E-05		-0.000113
		(9.24E-05)		(0.00015)
$M_{sum}\times$ `Money_topic`		-2.66E-05		4.13E-05
		(3.36E-05)		(5.46E-05)
Family	-0.02	-0.0221	-0.177***	-0.173***
	(0.0336)	(0.034)	(0.0546)	(0.0552)
Lords	0.0929*	0.0868	-0.00389	0.0278
	(0.0486)	(0.0569)	(0.0789)	(0.0925)
Clerks	0.041	0.0393	-0.00846	-0.00456
	(0.0456)	(0.046)	(0.0741)	(0.0747)
Workers	0.157**	0.158**	-0.033	-0.0354
	(0.0627)	(0.0631)	(0.102)	(0.103)
M_{sum}	8.29E-08		4.18E-07	
	(2.94E-07)		(4.78E-07)	
Observations	523	523	523	523
R^2	0.552	0.555	0.602	0.605

Note: Robust standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level. All models include Year Fixed Effects.