

Using Natural Language Processing to Classify Social Work Interventions

Abdulaziz Tijjani Bako, MPH; Heather L. Taylor, MPH, RDH; Kevin Wiley Jr, MPH; Jiaping Zheng, PhD; Heather Walter-McCabe, JD, MSW; Suranga N. Kasthurirathne, PhD; and Joshua R. Vest, PhD, MPH

Traditionally, efforts to improve health outcomes have primarily focused on medical care services.¹ However, consistent evidence suggests that social needs and risk factors have a more profound effect on individual and population health than medical care.²⁻⁵ Given their professional training and workflows, social workers within health care settings are uniquely positioned to deliver interventions to address patients' social needs.⁶ Moreover, social workers are increasingly becoming embedded in health care organizations to help address patients' social needs.⁷ However, tracking and evaluating interventions instituted by social workers remains challenging because social work (SW) interventions are largely documented as unstructured text data within electronic health records (EHRs).^{8,9} Unstructured documentation makes it difficult to systematically monitor and study SW interventions and services, as manual chart reviews are expensive and time-consuming and often require expert reviewers.¹⁰⁻¹² These limitations underscore the need for using novel information extraction methods, such as natural language processing (NLP) and machine learning (ML), to identify and classify interventions documented in unstructured EHR notes, such as SW notes.

Existing research has been successful in using NLP methods to identify and classify social needs with EHR data. Dorr et al¹³ demonstrated that NLP could be used to identify chronic stress, social isolation, financial insecurity, and housing insecurity in the clinical notes of primary care doctors. Conway et al¹⁴ developed the Moonstone system, which used rule-based classification to identify housing situations and social support within EHRs. Also, Cook et al¹⁵ used NLP and ML techniques to predict suicidal ideation from a text message intervention. These studies demonstrated the feasibility and applicability of using NLP approaches to detect social needs and risks in unstructured clinical data. Although a small number of studies have developed classification schemes for SW interventions using manual chart reviews, no study has sought to utilize NLP methods to identify and/or classify SW interventions.¹⁶⁻²¹

The purpose of this study was to extract and categorize SW interventions aimed at addressing patients' social needs by developing and applying NLP and ML algorithms. Development

ABSTRACT

OBJECTIVES: Health care organizations are increasingly employing social workers to address patients' social needs. However, social work (SW) activities in health care settings are largely captured as text data within electronic health records (EHRs), making measurement and analysis difficult. This study aims to extract and classify, from EHR notes, interventions intended to address patients' social needs using natural language processing (NLP) and machine learning (ML) algorithms.

STUDY DESIGN: Secondary data analysis of a longitudinal cohort.

METHODS: We extracted 815 SW encounter notes from the EHR system of a federally qualified health center. We reviewed the literature to derive a 10-category classification scheme for SW interventions. We applied NLP and ML algorithms to categorize the documented SW interventions in EHR notes according to the 10-category classification scheme.

RESULTS: Most of the SW notes (n = 598; 73.4%) contained at least 1 SW intervention. The most frequent interventions offered by social workers included care coordination (21.5%), education (21.0%), financial planning (18.5%), referral to community services and organizations (17.1%), and supportive counseling (15.3%). High-performing classification algorithms included the kernelized support vector machine (SVM) (accuracy, 0.97), logistic regression (accuracy, 0.96), linear SVM (accuracy, 0.95), and multinomial naive Bayes classifier (accuracy, 0.92).

CONCLUSIONS: NLP and ML can be utilized for automated identification and classification of SW interventions documented in EHRs. Health care administrators can leverage this automated approach to gain better insight into the most needed social interventions in the patient population served by their organizations. Such information can be applied in managerial decisions related to SW staffing, resource allocation, and patients' social needs.

Am J Manag Care. 2021;27(1):e24-e31. doi:10.37765/ajmc.2021.88580

of such classification tools supports health care organizations' assessment and measurement of a growing part of the nonmedical workforce. Additionally, this study highlights methodological approaches to examine SW interventions in situations in which only unstructured data are accessible. Health care organizations armed with data on SW interventions that have been instituted for their patient population will make more informed resource allocation, staffing, quality improvement, and program design decisions to address their patients' social needs. Moreover, automation of the classification of interventions will offer stakeholders an enhanced ability to quantify the specific impact of social workers' interventions on patients' health outcomes.

METHODS

We developed a classification scheme from the literature¹⁶⁻²¹ to categorize interventions instituted by social workers to address patients' social needs at Eskenazi Health, an urban safety-net health system. Eskenazi Health employs in-house social workers to address patients' social needs. These needs are identified through multiple pathways. For example, physicians and nurse practitioners may identify patients' social needs during consultations and refer patients to social workers. Additionally, social workers may review records of patients with scheduled appointments to identify social needs.

We used NLP and ML algorithms for automated categorization of SW interventions based on the classification scheme we developed. The [Figure](#) describes, in brief, the process and pipeline used for this classification (see [eAppendix A](#) [eAppendices available at [ajmc.com](#)]). This study was approved by the Indiana University Institutional Review Board.

Setting and Sample

We used patient record data from Eskenazi Health, a safety-net provider with a 300-bed hospital and a federally qualified health center serving the Indianapolis, Indiana, metropolitan area. The study sample included 408 patients with 815 SW encounters between October 1, 2016, and September 30, 2019.

Data

The study data were derived from Eskenazi Health's EHR (Epic). The EHR data included patients' clinical notes and their demographic and clinical characteristics. We obtained unstructured data containing free-text descriptions of the reason(s) for patient visits, the intervention(s) instituted by the social worker, and future visits or engagement plans from social workers' notes. We extracted SW notes, using an existing pipeline (nDepth²²), by searching clinical notes for the following search terms: "social work" or "MSW" or "LCSW", or "LSW", or "LMSW", or "LBSW." nDepth, an NLP tool

TAKEAWAY POINTS

Social work (SW) interventions are increasingly important in health care.

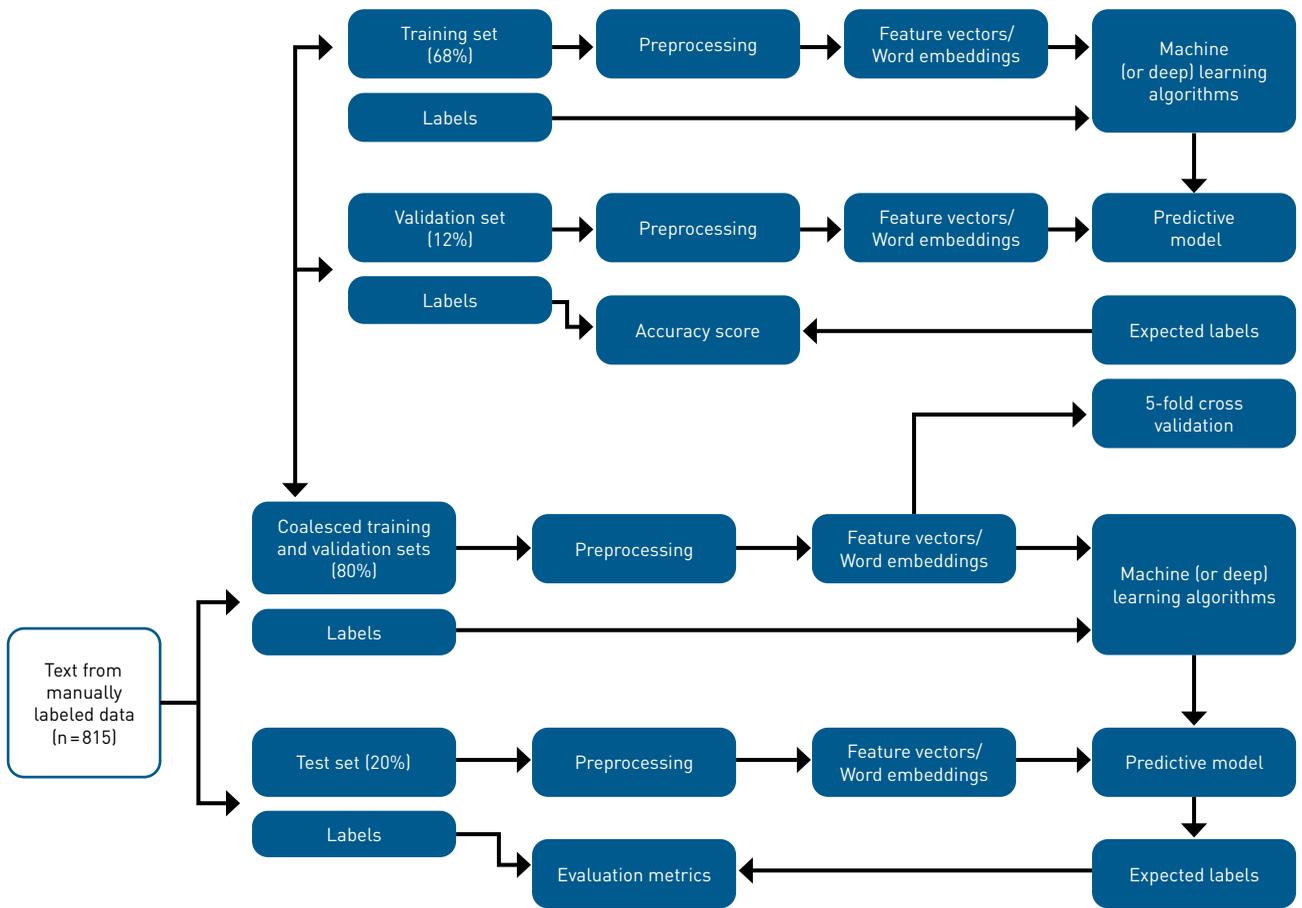
- ▶ Most of the SW notes examined in this study contained at least 1 SW intervention. In approximately 37% of SW notes, more than 1 intervention was identified, underscoring the need to address coexisting social needs among a sizable population.
- ▶ The most frequent SW interventions offered included care coordination, education, financial planning, referral to community services and organizations, and supportive counseling.
- ▶ Natural language processing and machine learning are inexpensive tools for automated identification and classification of SW interventions.
- ▶ This automated identification can support managerial decisions related to social worker staffing and allocation of resources to address patients' social needs.

developed by the Regenstrief Institute, conducts information retrieval and extraction from textual documents by adapting its existing pipeline to customizable search queries. nDepth has been used in other clinical domains, including gastrointestinal diseases and sarcopenia.^{23,24} For this study, nDepth was used to extract a sample of SW notes from the EHR. Our search with nDepth retrieved 1289 clinical notes. Three abstractors (A.T.B., H.L.T., K.W.) manually reviewed the retrieved notes and categorized them into notes written by a social worker and notes written by other professionals. Notes were deemed as written by a social worker if they were signed off by a social worker or if the wordings and structure of the notes clearly indicated documentation of activities by a social worker. The final sample for this study included 815 SW encounters with notes that were determined to be written by social workers during distinct SW encounters (see [eAppendix B](#) for details and examples).

Based on literature review and expert consultations, 3 members of the research team (A.T.B., H.L.T., K.W.) developed a classification scheme for SW interventions, with 10 non-mutually exclusive categories of SW interventions briefly defined as follows:

- **Financial planning:** resources for funding medications, treatment, or other care.²⁵
- **Supportive counseling:** emotional and mental health support services; addiction services.²⁶
- **Care coordination:** coordination of care continuity, case management, and arranging and coordinating home visits.²⁷
- **Education:** health/wellness programs, crisis intervention, and course planning.²⁸
- **Community service:** referrals to community, spiritual, and peer advocacy organizations; translational services; food assistance; clothing assistance.²⁹
- **Applications and reporting:** help with filing applications; filing mandatory reports such as those to Department of Child Services and Adult Protective Services.
- **Housing:** shelter provision; help with housing.
- **Transportation:** provision of transportation fares and passes; linkage with transportation services.

FIGURE. Classification Algorithm Pipeline



- **Durable medical equipment:** provision of medical equipment; repair, maintenance, and/or replacement of medical equipment.
- **Legal:** link to attorneys, law enforcement, and other legal services.

Inspired by the ConText algorithm,³⁰ 3 coders (A.T.B., H.L.T., K.W.) independently reviewed 100 randomly selected SW notes to identify verbs likely to indicate the presence of SW interventions in a sentence. The abstractors also identified contextual terms and negation terms associated with each verb, used to describe a social worker (social worker terms), as well as unique terms that map to each of the 10 intervention categories (see **Appendices C and D** for lexicon of terms). The abstractors manually coded the selected notes into their respective non-mutually exclusive SW intervention categories. We calculated the rate of agreement among the 3 coders using Fleiss’s kappa coefficients (see **Appendix E**).

Analysis

To categorize the notes based on our 10-category scheme, we used (1) rule-based classification algorithms using Python’s regular

expressions³¹; (2) ML algorithms such as the multinomial naive Bayes classification algorithm³² and multilabel (one vs rest, binary relevance, classifier chains, and label powerset)^{33,34} classification algorithms using logistic regression³³ and kernelized support vector machine (SVM) with radial basis function³⁵; and (3) a deep learning algorithm for multilabel classification: the long short-term memory (LSTM) recurrent neural network with a sigmoid activation function and binary cross-entropy loss function.³⁶ The LSTM model has a single input layer, an embedding layer, 1 LSTM layer with 128 nodes, and 1 output layer with 10 nodes, with each node in the output dense layer representing 1 of the SW intervention categories.

For the rule-based algorithm, we used Python’s³¹ in-built regular expressions to extract sentences that include a social worker term and one of the intervention verb terms, where the verb term is not preceded by one of the negative-lookbehind terms for intervention and not followed by one of the negative-lookahead terms. A review of SW notes and consultation with clinical social workers revealed that these sentences are likely to contain information about the type of intervention instituted. An SW note was deemed to indicate

the presence of an intervention category if the extracted sentences in the note also contain at least 1 of the intervention key terms for that category.

For the implementation of the ML and deep learning algorithms, we randomly divided the data into training (68%), validation (12%), and test (20%) sets (see Figure). We preprocessed the text in each of these data sets as follows: We tokenized sentences in each SW note into individual words (tokens), removed stop words (eg, 1- and 2-letter words, compositions, prepositions), applied part of speech tagging to the tokens, and stemmed the tokens down to their root words. Next, for implementation of the ML algorithms, we created feature vectors using the term frequency-inverse document frequency vectorizer with single words (unigrams), 2 consecutive words (bigrams), and 3 consecutive words (trigrams). Relatedly, for the LSTM model, we created word embeddings using the GloVe word embeddings³⁷ to convert text inputs to their numeric counterparts.

We used the preprocessed feature vectors (or word embeddings in case of the LSTM model) from the training data set to initially train each classifier and the validation data set to test the accuracy of the trained classifier. We implemented multiple iterations of training and validation for each classifier while tuning and optimizing the hyperparameters of the classification algorithm to attain higher accuracy. Upon achieving the highest accuracy score possible for each classifier, we coalesced the training and validation sets and trained a final classifier on them using the optimal hyperparameters. We also used 5-fold cross-validation to evaluate the performance of the logistic regression, kernelized SVM, linear SVM, and multinomial naive Bayes algorithms on the full training data (see [eAppendix F \[A and B\]](#)). Finally, for each intervention category, we evaluated the performance of the rule-based, logistic regression, kernelized SVM, linear SVM, and multinomial naive Bayes algorithms on test data using accuracy; precision or positive predictive value (PPV); recall (sensitivity); F1-score, which is the harmonic mean of precision (PPV) and recall (sensitivity); specificity; and area under the curve (AUC).

RESULTS

Our final sample included 408 patients with 815 SW encounters. Descriptive information about the sample and categories of social worker interventions is available in [eAppendix G](#). Briefly, 43% of the patients were Hispanic; their mean age was 38.7 years; and the majority (64%) were women.

Of the 815 SW encounters, almost three-fourths ($n = 598$; 73.4%) contained at least 1 SW intervention. More specifically, 217 (26.6%) did not include any description of a SW intervention, 295 (36.2%) included 1 SW intervention, 207 (25.4%) included 2 interventions, and 96 (11.8%) included 3 or more interventions. The highest number of interventions in a single SW note was 6, which was observed in only 5 of the 815 encounters. The most common SW interventions in the notes included care coordination (21.5%), education (21.0%), financial planning (18.5%), referral to community services and organizations (17.1%), and supportive counseling (15.3%) ([Table 1](#)).

Models with the highest accuracy included multilabel (one vs rest) classifier with kernelized SVM (accuracy, 0.97), multilabel (one vs rest) classifier with logistic regression (accuracy, 0.96), linear SVC (accuracy, 0.95), and multinomial naive Bayes classifier (accuracy, 0.92) (see [eAppendix H](#)). Precision (PPV) score was generally higher than the recall (sensitivity) score in the kernelized SVM, logistic regression, and linear SVM for most of the intervention categories ([Table 2 \[A and B\]](#)). However, for the multinomial naive Bayes classifier, the recall score was higher than the sensitivity score in the financial planning, care coordination, community service, and education intervention categories. Moreover, the multinomial naive Bayes classifier offered the highest recall scores for the financial planning (recall, 0.86) and community service (recall, 0.92) categories ([Table 2 \[B\]](#)). Linear SVM provided the best evaluation metrics for the financial planning category (accuracy, 0.96; precision, 0.91, recall, 0.91; F1-score, 0.91; specificity, 0.98; AUC, 0.94). The F1-scores for the logistic regression and kernelized SVM algorithms in the care coordination, community services, and transportation categories were high (0.82-0.94). The multinomial naive Bayes algorithms had the worst F1-score (0.00) in the legal intervention category. Linear SVM offered the best evaluation metrics for the housing (accuracy, 1.00; precision, 1.00; recall, 1.00; F1-score, 1.00; specificity, 1.00; AUC, 1.00) and durable medical equipment (accuracy, 0.99; precision, 0.80; recall, 1.00; F1-score, 0.89; specificity, 0.99; AUC, 0.90) categories ([Table 2 \[A-C\]](#)).

DISCUSSION

This study described the extent and nature of SW activities in primary care using NLP algorithms. These findings illustrate the key roles that social workers play in connecting patients to other aspects of the health care system, brokering connections with non-health care-related organizations, and providing patient education.

As brokers, social workers link patients to critically needed resources and services.^{38,39} This broker function was evident in our study, in which care coordination interventions and referrals to community-based organizations were frequently offered. SW interventions require brokering services for patients among different organizations/providers, maintaining communication, and identifying needed resources from referral partners. These efforts may lead to improved clinical outcomes. For instance, Bronstein et al⁴⁰ found that social worker involvement in postdischarge care coordination, which involved directing patients to organizations that address food insecurity, clothing, and peer advocacy needs, was associated with fewer 30-day hospital readmissions. Pruitt et al⁴¹ showed that referrals to community-based organizations offered by social workers were associated with lower health care costs. Our study provided evidence that social workers in primary care offer interventions known to aid health care organizations' efforts to improve quality of care, reduce costs, and effectively manage population health. We intend to investigate the value

[\(CONTINUED ON PAGE E30\)](#)

TABLE 1. Social Interventions Offered by Social Workers

| Interventions | Frequency (%) (N = 815) | Examples of social work note |
|----------------------------|----------------------------|---|
| Financial planning | 151 (18.5) | <ul style="list-style-type: none"> • SW provided a medication voucher for patient for the needed medications. • SW provided \$100 in TCC gift cards. • SW submitted a new contingency form and verified with the pharmacy that mom had received patient’s medication. • MSW also obtained approval from SW leadership to assist with the cost of the patient’s prescriptions. |
| Care coordination | 175 (21.5) | <ul style="list-style-type: none"> • SW will follow up with the neuropsychiatric department to ensure the patient is scheduled for an evaluation. • SW met with physician to discuss case. • SW submitted home health referral. • SW will continue to follow the patient and family toward developing an appropriate discharge plan. • SW completed and faxed a referral for a public health nurse home visit. • SW has sent message to provider to see if she will proceed to order the rehab and if it is to be outpatient or home-based. • SW sent fax referral for patient to be evaluated by neuropsychiatry as well. |
| Community service | 139 (17.1) | <ul style="list-style-type: none"> • Completed referral to safe sleep program in order for patient to get a crib for baby. • Brief home visit today to provide Christmas gifts to patient and family, donated by GRACE team. • SW referred the family to NACS for assistance. • [SW] placed online referral to CICOA to see if they are able to assist. |
| Education | 171 (21.0) | <ul style="list-style-type: none"> • SW provided [the] mother with resource information for the Down Syndrome Indiana group. • Social worker informed patient of the team’s recommendation for SAR. • SW expressed understanding and educated mom and dad on the benefits of CSHCS as supplemental insurance. • Patient was provided a list of detox/in/outpatient services. • Social worker educated patient on the therapy services hospice agency offers to deal with grief and loss. |
| Supportive counseling | 125 (15.3) | <ul style="list-style-type: none"> • SW offered counseling and problem-solving to assist and patient agreed. • SW will provide individual therapy to address patient’s active symptoms of depression by exploring cognitive distortions through [CBT] and creating self-soothing techniques. • Social worker provided behavioral management of pain. • Patient seen for CBT for depressed mood. • SW offered grief counseling and problem-solving techniques and patient agreed. • SW processed coping strategies as well as discussed mother’s supports. |
| Applications and reporting | 101 (12.4) | <ul style="list-style-type: none"> • SW placed call to FSSA to follow up on patient’s application. • SW assisted mother with filing DCS report. • SW will also follow up with APS. • SW filed a [form] 310 for communication with DCS. • Social worker phoned APS to inquire about the status of patient’s APS investigation. |
| Housing | 39 (4.8) | <ul style="list-style-type: none"> • A cot was secured [for patient] at Wheeler Mission for Women. • SW followed up on lodging accommodations which are serving [patient] and brother. • Social work discussed housing options with patient. • Provided patient with a subsidized housing list. • SW provided lodging and parking to patient. |
| Transportation | 90 (11.0) | <ul style="list-style-type: none"> • [SW] assisted patient with finding the appropriate route for transportation through Indygo. • Bus pass provided [by SW]. • Patient was provided an all-day bus pass for transportation home at discharge. • [SW] authorized yellow cab voucher to take patient to hospital. • SW discussed and received approval from manager to use airline-donated tickets for the family to utilize. |
| DME | 37 (4.5) | <ul style="list-style-type: none"> • Discussed with MD and received DME order for hospital bed. • SW is currently working on a hospital bed for patient. • Social worker faxed order, facesheet, and progress notes to Community DME for wheelchair assistance. • SW confirmed with the daughter that the raised toilet seat did get delivered. • SW received the order for patient’s wheelchair cushions. |
| Legal | 26 (3.2) | <ul style="list-style-type: none"> • SW provided mother with contact and resource information for Legal Aid. • SW encouraged patient to follow up with police. • SW referred patient to medical legal services to see if this is able to be expunged. • SW referred patients to Medical-Legal Partnership to evaluate the possibility of changing patient’s immigration status. • SW left a message for patient’s case manager to call back to reinforce need for [patient] to receive meds while in jail. |

APS, Adult Protective Services; CBT, cognitive-behavioral therapy; CICOA, CICOA Aging & In-Home Solutions; CSHCS, Children’s Special Health Care Services; DCS, Department of Child Services; DME, durable medical equipment; FSSA, Family and Social Service Administration; MSW, medical social worker; NACS, Neighborhood Alliance for Child Safety; SAR, subacute rehabilitation; SW, social work(er); TCC, The Customer Connection.

TABLE 2. Evaluation Metrics for Classification Algorithms

| A. Evaluation metrics for logistic regression and kernelized SVM algorithms | | | | | | | | | | | | |
|---|-------------------------------|-----------------|----------------------|----------|-------------|------|----------------|-----------------|----------------------|----------|-------------|------|
| Performance metrics | Logistic regression | | | | | | Kernelized SVM | | | | | |
| | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | Specificity | AUC | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | Specificity | AUC |
| Financial planning | 0.95 | 1.0 | 0.73 | 0.85 | 1.00 | 0.87 | 0.95 | 1.00 | 0.73 | 0.85 | 1.00 | 0.87 |
| Care coordination | 0.93 | 0.89 | 0.89 | 0.89 | 0.98 | 0.93 | 0.96 | 0.89 | 0.89 | 0.89 | 0.98 | 0.94 |
| Community service | 0.96 | 0.85 | 0.81 | 0.83 | 0.98 | 0.89 | 0.96 | 0.89 | 0.76 | 0.82 | 0.99 | 0.87 |
| Education | 0.93 | 1.0 | 0.63 | 0.77 | 1.00 | 0.81 | 0.96 | 1.00 | 0.74 | 0.85 | 1.00 | 0.87 |
| Supportive counseling | 0.98 | 1.0 | 0.89 | 0.94 | 1.00 | 0.94 | 0.98 | 1.00 | 0.89 | 0.94 | 1.00 | 0.94 |
| Filing applications and reporting | 0.95 | 1.0 | 0.55 | 0.71 | 1.00 | 0.78 | 0.95 | 1.00 | 0.56 | 0.71 | 1.00 | 0.78 |
| Housing | 0.98 | 1.0 | 0.71 | 0.83 | 1.00 | 0.85 | 0.99 | 1.00 | 0.71 | 0.83 | 1.00 | 0.86 |
| Transportation | 0.96 | 1.0 | 0.70 | 0.82 | 1.00 | 0.84 | 0.96 | 1.00 | 0.70 | 0.82 | 1.00 | 0.85 |
| DME | 0.99 | 1.0 | 0.66 | 0.8 | 1.00 | 0.83 | 0.99 | 0.67 | 0.67 | 0.67 | 0.99 | 0.83 |
| Legal | 0.99 | 1.0 | 0.50 | 0.67 | 1.00 | 0.75 | 0.99 | 1.00 | 0.50 | 0.67 | 1.00 | 0.75 |
| B. Evaluation metrics for multinomial naive Bayes and linear SVM algorithms | | | | | | | | | | | | |
| Performance metrics | Multinomial naive Bayes model | | | | | | Linear SVM | | | | | |
| | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | Specificity | AUC | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | Specificity | AUC |
| Financial planning | 0.91 | 0.75 | 0.86 | 0.80 | 0.92 | 0.85 | 0.96 | 0.91 | 0.91 | 0.91 | 0.98 | 0.94 |
| Care coordination | 0.85 | 0.52 | 0.85 | 0.65 | 0.85 | 0.75 | 0.90 | 0.76 | 0.50 | 0.60 | 0.97 | 0.84 |
| Community service | 0.89 | 0.58 | 0.92 | 0.71 | 0.88 | 0.78 | 0.94 | 0.79 | 0.79 | 0.79 | 0.96 | 0.88 |
| Education | 0.87 | 0.65 | 0.71 | 0.68 | 0.91 | 0.79 | 0.88 | 0.70 | 0.68 | 0.69 | 0.93 | 0.81 |
| Supportive counseling | 0.91 | 0.89 | 0.68 | 0.77 | 0.99 | 0.92 | 0.93 | 0.88 | 0.60 | 0.71 | 0.99 | 0.91 |
| Filing applications and reporting | 0.89 | 0.57 | 0.57 | 0.57 | 0.94 | 0.75 | 0.94 | 0.89 | 0.89 | 0.89 | 0.75 | 0.88 |
| Housing | 0.98 | 0.75 | 0.60 | 0.67 | 0.99 | 0.87 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Transportation | 0.95 | 0.92 | 0.63 | 0.75 | 0.99 | 0.94 | 0.97 | 0.92 | 0.79 | 0.86 | 0.99 | 0.96 |
| DME | 0.96 | 0.67 | 0.25 | 0.36 | 0.99 | 0.81 | 0.99 | 0.80 | 1.00 | 0.89 | 0.99 | 0.90 |
| Legal | 0.95 | 0.00 | 0.00 | 0.00 | 1.00 | 0.95 | 0.96 | 0.75 | 0.38 | 0.50 | 0.99 | 0.85 |
| C. Evaluation metrics for rule-based algorithm | | | | | | | | | | | | |
| Performance metrics | Rule-based | | | | | | | | | | | |
| | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | | | | | | | | |
| Financial planning | 0.84 | 0.56 | 0.56 | 0.56 | | | | | | | | |
| Care coordination | 0.87 | 0.56 | 0.56 | 0.56 | | | | | | | | |
| Community service | 0.94 | 0.66 | 0.66 | 0.66 | | | | | | | | |
| Education | 0.81 | 0.48 | 0.48 | 0.48 | | | | | | | | |
| Supportive counseling | 0.52 | 0.27 | 0.27 | 0.27 | | | | | | | | |
| Filing applications and reporting | 0.96 | 0.61 | 0.61 | 0.61 | | | | | | | | |
| Housing | 0.96 | 0.63 | 0.63 | 0.63 | | | | | | | | |
| Transportation | 0.82 | 0.43 | 0.43 | 0.43 | | | | | | | | |
| DME | 0.99 | 0.71 | 0.71 | 0.71 | | | | | | | | |
| Legal | 0.95 | 0.32 | 0.32 | 0.32 | | | | | | | | |

AUC, area under the curve; DME, durable medical equipment; PPV, positive predictive value; SVM, support vector machine.

(CONTINUED FROM PAGE E27)

of SW interventions toward improving quality outcomes in our future studies.

In our study, social workers educated patients, in 21% of encounters, about the availability of resources and interpretation of medical information. Patient education contributes significantly toward better health outcomes.⁴² The role of social workers as providers of patient education thus highlights the significance of social workers in primary care teams. Also, financial planning, which involves assisting patients with medication, insurance, and benefits, was a common SW intervention in this study. Similar to previous studies,^{38,43} our study found that social workers primarily address the impact of finances on health care access by providing assistance with medication, insurance, and benefits, rather than directly providing monetary assistance.

Social workers are the largest group of providers of mental health services in the United States.⁴⁴ In Indiana, social workers can offer a wide range of mental health services, including counseling and treatment for substance use, provided that they obtain a state license to practice mental health.⁴⁵ However, some authors have opined that social workers employed in health care settings may not have the time to directly provide mental health services, given the need to frequently coordinate and allocate resources for these patients.^{38,46} For example, in a study on patients with diabetes, Rabovsky et al³⁸ found that social workers did not directly address patients' mental health issues. Rather, they referred patients to mental health providers. The practice pattern and/or expertise of the social workers in Rabovsky et al³⁸ may explain why social workers outsourced counseling interventions to external organizations rather than providing the counseling services themselves. In our study, social workers addressed mental and behavioral health issues by directly providing supportive counseling interventions to patients. Thus, it is conceivable, given our findings, that social workers can multitask and interchange roles, thereby providing patients in health care settings with direct counseling and tangible resources or referrals. Finally, in previous studies, the nature of the advocacy role of social workers was not evident and underexplored.^{38,39,47} However, in our study, we found that social workers helped patients in completing and filing insurance coverage or housing applications and in reporting child and adult domestic issues to the appropriate social agencies. This highlights the role of social workers as patient advocates, particularly for those most vulnerable.

Similar to the findings of Pooler et al,⁴⁷ we found that approximately two-fifths of the SW notes included more than 1 SW intervention, indicating that a substantial number of patients have co-occurring social needs.⁴⁷ The presence of a sizable proportion of patients receiving multiple social interventions indicates the need for effective SW staffing and differential use of more experienced social workers to manage patients with co-occurring social needs.

Although unstructured free-text narratives within EHRs are conducive to NLP- and ML-based classification methods, previous research suggests that social needs interventions are rarely explored

using these methods.^{11,16-21} Our study shows that NLP and ML can effectively be used to determine the types of SW interventions suggested by social workers operating within health care systems. Consistent with the findings of previous studies, ML- and deep learning-based classification algorithms performed better than rule-based classification methods.^{48,49} However, the multilabel LSTM model did not perform as well as some ML approaches. This finding may be due to the small sample size in this study, as deep learning algorithms are known to perform poorly when the sample size is small.⁵⁰

Limitations

This study has several limitations. First, although we derived our SW intervention categories from consultation with experts and peer-reviewed literature, our classification scheme may not be exhaustive. Second, the small nature of our sample may limit the performance of our classification algorithms on new test data. However, for most of the intervention categories our evaluation metrics are satisfactory. In addition, our models were trained using data from a single health system, which weakens the generalizability of our findings to other hospital systems or other diverse populations. Lastly, the range of intervention categories and their relative frequencies in this study may be a reflection of the characteristics of the population under study and/or the practice pattern of the social workers in our study site, rather than the general primary care population.

CONCLUSIONS

Contextual details of interventions instituted by social workers, which are available in EHR notes, highlight how social needs are addressed by social workers. NLP and ML can be utilized for automated identification and classification of SW interventions documented in EHRs. These methods have the advantage of being scalable and can be automated and integrated into EHR systems. Thus, NLP and ML can be leveraged by health care administrators to gain better insight into the most needed social interventions in their patient populations, thereby helping organizations make better decisions related to SW staffing, resource allocation, and patients' social needs. ■

Acknowledgments

The authors thank the Robert Wood Johnson Foundation for providing the financial support necessary for conducting this research through their Systems for Action National Coordinating Center (ID 73485). They also thank Data Core Services of the Regenstrief Institute for providing them with the data used in this study.

Research reported in this publication/press release was supported by the National Library of Medicine of the National Institutes of Health under award No. T15LM012502. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Library of Medicine.

Author Affiliations: Department of Health Policy & Management, Indiana University Richard M. Fairbanks School of Public Health (ATB, HLT, KW, JRV), Indianapolis, IN; Center for Outcomes Research, Houston Methodist Research Institute (ATB), Houston, TX; Center for Biomedical Informatics, Regenstrief Institute (HLT, KW, SNK, JRV), Indianapolis, IN; School of Informatics and

Computing, Indiana University–Purdue University Indianapolis (JZ), Indianapolis, IN; Wayne State University School of Social Work (HW-M), Detroit, MI; Wayne State University School of Law (HW-M), Detroit, MI; Department of Pediatrics, Indiana University School of Medicine (SNK), Indianapolis, IN.

Source of Funding: Robert Wood Johnson Foundation (ID 73485) and National Library of Medicine of the National Institutes of Health (T15LM012502).

Author Disclosures: Dr Kasthurirathne is a founder of Upstrooms, LLC, and reports patents pending and stock ownership. Dr Vest is a founder and board member of Upstrooms, LLC. The remaining authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (ATB, JZ, SNK, JRV); acquisition of data (HLT, SNK, JRV); analysis and interpretation of data (ATB, HLT, KW, HW-M, SNK); drafting of the manuscript (ATB, KW, SNK); critical revision of the manuscript for important intellectual content (ATB, HLT, KW, JZ, SNK, JRV); statistical analysis (ATB); provision of patients or study materials (JRV); administrative, technical, or logistic support (HW-M, SNK, JRV); and supervision (JZ, HW-M, SNK).

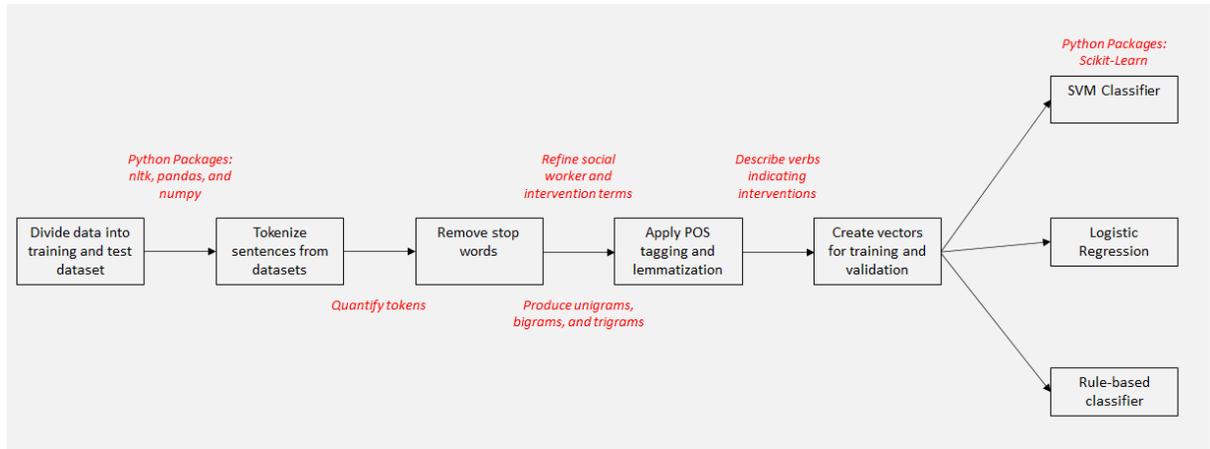
Address Correspondence to: Abdalaziz Tijjani Bako, MPH, Department of Health Policy and Management, Richard M. Fairbanks School of Public Health, Indiana University–Purdue University Indianapolis, 1050 Wishard Blvd, Indianapolis, IN 46202. Email: atbako@iu.edu.

REFERENCES

- Heiman HJ, Artiga S. Beyond health care: the role of social determinants in promoting health and health equity. *Kaiser Family Foundation*. November 2015. Accessed April 24, 2020. <https://www.issuelab.org/resources/22899/22899.pdf>
- WHO Commission on Social Determinants of Health, World Health Organization. *Closing the Gap in a Generation: Health Equity Through Action on the Social Determinants of Health*. World Health Organization; 2008.
- Seabrook JA, Avison WR. Socioeconomic status and cumulative disadvantage processes across the life course: implications for health outcomes. *Can Rev Sociol*. 2012;49(1):50-68. doi:10.1111/j.1755-618x.2011.01280.x
- Marmot MG, Shipley MJ, Rose G. Inequalities in death—specific explanations of a general pattern? *Lancet*. 1984;323(8384):1003-1006. doi:10.1016/s0140-6736(84)92337-7
- Galea S, Tracy M, Hoggatt KJ, Dimaggio C, Karpati A. Estimated deaths attributable to social factors in the United States. *Am J Public Health*. 2011;101(8):1456-1465. doi:10.2105/AJPH.2010.300086
- Ruth BJ, Wachman MK, Marshall JW, et al. Health in all social work programs: findings from a US national analysis. *Am J Public Health*. 2017;107(S3):S267-S273. doi:10.2105/AJPH.2017.304034
- Occupational employment statistics. U.S. Bureau of Labor Statistics. March 31, 2020. Accessed May 19, 2020. <https://www.bls.gov/oes/tables.htm>
- Vest JR, Grannis SJ, Haut DP, Halverson PK, Menachemi N. Using structured and unstructured data to identify patients' need for services that address the social determinants of health. *Int J Med Inform*. 2017;107:101-106. doi:10.1016/j.ijmedinf.2017.09.008
- National Association of Social Workers, Ohio Chapter. Clinical documentation. National Association of Social Workers. Accessed November 6, 2020. https://cdn.yamaws.com/www.naswoh.org/resource/resmgr/Private_Practice/Clinical_Documentation_NASW_.pdf
- Kim EJ, Abrahams S, Uwemedeo O, Conigliaro J. Prevalence of social determinants of health and associations of social needs among United States adults, 2011-2014. *J Gen Intern Med*. 2020;35(5):1608-1609. doi:10.1007/s11606-019-05362-3
- Hatef E, Rouhizadeh M, Tia I, et al. Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system. *JMIR Med Inform*. 2019;7(3):e13802. doi:10.2196/13802
- Kusnoro SV, Koonce TY, Hurler ST, et al. Collection of social determinants of health in the community clinic setting: a cross-sectional study. *BMC Public Health*. 2018;18(1):550. doi:10.1186/s12889-018-5453-2
- Dorr D, Bejan CA, Pizzimenti C, Singh S, Storer M, Quinones A. Identifying patients with significant problems related to social determinants of health with natural language processing. *Stud Health Technol Inform*. 2019;264:1456-1457. doi:10.3233/SHTI190482
- Conway M, Keyhani S, Christensen L, et al. Moonstone: a novel natural language processing system for inferring social risk from clinical narratives. *J Biomed Semantics*. 2019;10(1):6. doi:10.1186/s13326-019-0198-0
- Cook BL, Progovac AM, Chen P, Mullin B, Hou S, Baca-Garcia E. Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. *Comput Math Methods Med*. 2016;2016:8708434. doi:10.1155/2016/8708434
- Abeyta N, Freeman ES, Primack D, et al. SCIRehab Project series: the social work/case management taxonomy. *J Spinal Cord Med*. 2009;32(3):336-342. doi:10.1080/10790268.2009.11760787
- Pockett R, Lord B, Dennis J. The development of an Australian national classification system for social work practice in health care. *Soc Work Health Care*. 2001;34(1-2):177-193. doi:10.1080/00981380109517025
- Anderson RL, Estle G. Predicting level of mental health care among children served in a delivery system in a rural state. *J Rural Health*. 2001;17(3):259-265. doi:10.1111/j.1748-0361.2001.tb00963.x
- Viron M, Bello I, Freudenreich O, Shtasel D. Characteristics of homeless adults with serious mental illness served by a state mental health transitional shelter. *Community Ment Health J*. 2014;50(5):560-565. doi:10.1007/s10597-013-9607-5
- Rosen A, Proctor EK, Staudt M. Targets of change and interventions in social work: an empirically based prototype for developing practice guidelines. *Res Soc Work Pract*. 2003;13(2):208-233. doi:10.1177/1049731502250496
- Hammond FM, Gassaway J, Abeyta N, Freeman ES, Primack D. The SCIRehab project: social work and case management. social work and case management treatment time during inpatient spinal cord injury rehabilitation. *J Spinal Cord Med*. 2011;34(2):216-226. doi:10.1179/107902611X12971826988291
- nDepth. Regenstrief Institute. Accessed September 22, 2019. <https://www.regenstrief.org/implementation/ndepth/>
- Imler TD, Ring N, Crabb DW. Medical, social, and legal risks to predict alcoholic liver disease using natural language processing and advanced analytics. *Gastroenterology*. 2015;148(4, suppl 1):S499. doi:10.1016/S0016-5085(15)31674-7
- Moorthi RN, Liu Z, El-Azab SA, et al. Sarcopenia, frailty and cachexia patients detected in a multisystem electronic health record database. *BMC Musculoskelet Disord*. 2020;21(1):508. doi:10.1186/s12891-020-03522-9
- Social workers in hospitals & medical centers: occupational profile. National Association of Social Workers. 2011. Accessed April 24, 2020. <https://www.socialworkers.org/LinkClick.aspx?fileticket=070d1XW1R2w%3D&portalid=0>
- Clinical social work: your guide in 2020. Online MSW Programs. Accessed December 2, 2020. <https://www.onlinemswprograms.com/careers/types-of-social-work/guide-to-clinical-social-work.html>
- What is discharge planning for social workers? *Social Solutions*. May 27, 2013. Accessed December 18, 2019. <https://www.socialsolutions.com/blog/what-is-discharge-planning-for-social-workers/>
- Role of school social worker. School Social Work Association of America. Accessed December 18, 2019. <https://www.sswaa.org/school-social-work>
- What types of community services are available? *PediatricEducation.org*. November 16, 2009. Accessed December 18, 2019. <https://pediatriceducation.org/2009/11/16/what-types-of-community-services-are-available/>
- Harkema H, Dowling JN, Thornblade T, Chapman WW. ConText: an algorithm for determining negation, experience, and temporal status from clinical reports. *J Biomed Inform*. 2009;42(5):839-851. doi:10.1016/j.jbi.2009.05.002
- van Rossum G Jr, Drake FL. *The Python Language Reference Manual*. Network Theory; 2011.
- Kibriya AM, Frank E, Pfahringer B, Holmes G. Multinomial naive Bayes for text categorization revisited. In: *AI 2004: Advances in Artificial Intelligence*; 2004:488-499. doi:10.1007/978-3-540-30549-1_43
- Bishop CM. *Pattern Recognition and Machine Learning*. Springer; 2016.
- Read J, Pfahringer B, Holmes G, Frank E. Classifier chains for multi-label classification. *Mach Learn*. 2011;85(3):333-359. doi:10.1007/s10994-011-5256-5
- Smola AJ, Schölkopf B. A tutorial on support vector regression. *Statistics and Computing*. 2004;14(3):199-222. doi:10.1023/b:stco.0000035301.49549.88
- Sundermeyer M, Ney H, Schlüter R. From feedforward to recurrent LSTM neural networks for language modeling. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2015;23(3):517-529. doi:10.1109/taslp.2015.2400218
- Pennington J, Socher R, Manning C. GloVe: global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*; 2014. doi:10.3115/v1/d14-1162
- Rabovsky AJ, Rothberg MB, Rose SL, Brateanu A, Kou L, Misra-Hebert AD. Content and outcomes of social work consultation for patients with diabetes in primary care. *J Am Board Fam Med*. 2017;30(1):35-43. doi:10.3122/jabfm.2017.01.160177
- Craig SL, Muskat B. Bouncers, brokers, and glue: the self-described roles of social workers in urban hospitals. *Health Soc Work*. 2013;38(1):7-16. doi:10.1093/hsw/hls064
- Bronstein LR, Gould P, Berkowitz SA, James GD, Marks K. Impact of a social work care coordination intervention on hospital readmission: a randomized controlled trial. *Soc Work*. 2015;60(3):248-255. doi:10.1093/swj/swv016
- Pruitt Z, Emechebe N, Quast T, Taylor P, Bryant K. Expenditure reductions associated with a social service referral program. *Popul Health Manag*. 2018;21(6):469-476. doi:10.1089/pop.2017.01199
- Ramsubeik K, Ramrattan LA, Kaeley GS, Singh JA. Effectiveness of healthcare educational and behavioral interventions to improve gout outcomes: a systematic review and meta-analysis. *Ther Adv Musculoskelet Dis*. 2018;10(12):235-252. doi:10.1177/1759720X18807117
- Craig SL, Bejan R, Muskat B. Making the invisible visible: are health social workers addressing the social determinants of health? *Soc Work Health Care*. 2013;52(4):311-331. doi:10.1080/00981389.2013.764379
- Dorn C. Public stakeholder listening session on strategies for improving parity for mental health and substance use disorder coverage. HHS. July 27, 2017. Accessed August 18, 2020. <https://www.hhs.gov/programs/topic-sites/mental-health-parity/achieving-parity/cures-act-parity-listening-session/comments/behavioral-health-care-providers/national-association-of-social-workers/index.html>
- Indiana licensure information. National Association of Social Workers Indiana Chapter. Accessed August 18, 2020. <https://www.naswin.org/page/IndianaLicensure>
- Craig SL, Betancourt L, Muskat B. Thinking big, supporting families and enabling coping: the value of social work in patient and family centered health care. *Soc Work Health Care*. 2015;54(5):422-443. doi:10.1080/00981389.2015.1017074
- Pooler J, Liu S, Roberts A. Older adults and unmet social needs: prevalence and health implications. IMPAQ International. November 2017. Accessed March 3, 2020. https://www.impaqint.com/sites/default/files/files/SDOH%20among%20older%20adults%202017_issueBrief.pdf
- Hartmann J, Huppertz J, Schamp C, Heitmann M. Comparing automated text classification methods. *Int J Res Mark*. 2019;36(1):20-38. doi:10.1016/j.ijresmar.2018.09.009
- Atmadja AR, Purwarianti A. Comparison on the rule based method and statistical based method on emotion classification for Indonesian Twitter text. *2015 International Conference on Information Technology Systems and Innovation (ICITSI)*. 2015. doi:10.1109/icitsi.2015.7437692
- Oleyunik M, Kugic A, Kasáč Z, Kreuzthaler M. Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification. *J Am Med Inform Assoc*. 2019;26(11):1247-1254. doi:10.1093/jamia/ocz149

Visit ajmc.com/link/88580 to download PDF and eAppendix

eAppendix A. Social needs intervention text preprocessing



eAppendix B. Key terms for intervention categories used in rule-based algorithm

Of the 815 notes determined to be social work notes, 735 were duly signed off by a social worker. We determined that the remaining 80 notes were written by the social worker because the wordings and structure of these notes clearly indicate documentation of activities by a social worker, and thus, conform with the duly signed notes. The wordings we looked for included phrases like “This social worker met with patient...,” “social work assessment: (at the beginning of the note),” and the like. Also, if the structure of the unsigned note is that of a standard social work assessment, with the writer of the notes mainly identifying social needs and/or documenting social work interventions, we decided that the note is written by a social worker. Sections of a standard social work assessment include: Patient demographics, diagnosis, biopsychosocial assessment, health literacy assessment, living arrangements, social support, activities of daily living and functional status, socioeconomic needs, income source, transportation, legal information, and risk of abuse. Of note, we further reviewed the notes with HM, who is an expert in the field of social work, to resolve coding ambiguities.

eAppendix C. Key terms for intervention categories used in rule-based algorithm

| Social intervention Categories | Key terms |
|---------------------------------------|---|
| Financial Planning | anthem, insurance, medicare, advantage, financial counsel, sliding scale, self-pay, self pay, selfpay, medicaid(not medicaid cab), coverage, ssi , money, pay, free of charge, foc , HIP, ssdi, entitlements, united health, supplemental security, fund, worker(s) comp, christmas assistance, gas assistance, gas card, power account, cost, fee , waived, TANF, IMPACT, walmart card, employment, job, gift spend, CSHCS, FSSA, cash, ADAP |
| Care coordination | Physical therapy, PT, home health, hh, hha, hhc, care coordination, hha, plan (not followed by “:”), assisted living, nursing home, long-term care, long term care, longterm care, day care,daycare, personal care, home visit, nurse visit, referral, end of life, sleep, palliative, study, schedule appointment, referral, transition, transitional, alternative living, medication, outpatient therapy, linking to a pcp, transition, SAR, VNS |
| Supportive counseling and services | Psych, bereavement, crisis intervention, alcohol, substance abuse, culture, coping, adjustment, comorbid, chaplain, sexualit, rehabilitation, counsel(not financial counsel), behavioral management, problem-solving, depression, individual therapy, behavior, cogniti, soothing, relaxation CD, supportive, emotional support, couples counseling, supportive listening. |
| Education | class, information, educat, school, GED, guardianship, transition checklist, provided resource, list of facilit, facilit + list, explain, discuss, encourage, literacy, pamphlet, answer questions, EIP, handout, |
| Community/In-house services | food, food pantries, clothing meal(s), home podiatry, SNAP, M.O.wheels (meals on wheels), Caregivers, BABES coupon, voucher, open-door, care coordination, vocation, lifeline, department of aging, obama phone, school supplies, shoes, nacs, RMH, first steps,cshsc, chw, NCCS, Julian Center, Ronald McDonald house, Associations, organizations, support group, living center, senior care, creative change program, Healthy families, community action, group, St. Vincent DePaul, dip-in, ALS, GRACE team, Brooke’s Place, safe sleep program, YMCA,CICOA |
| Applications and Reporting | POST form, ROI form, Family authorization form, Release of information, advanced directive, FMLA, documentation/identification, pre-cert, birth certificate, IPA |

| | |
|-----------|---|
| | Agreement, Client Conduct Statement, and Duty to Warm Statement, child support report, adult protective service report, dcs, aps, CW |
| Housing | Lodging, housing, parking, feeling safe (in-home), safe at home, electricity, running water, power, Homeless Initiative Program, pedigo, shelter, HUD, accomodation, florence house |
| Transport | medicaid cab, cab, indigo, bus, taxi, transport, travel arrangement, shuttle |
| DME | medical equipment, cane, walker, manual chair, wheelchair, electric chair, shower chair, test strips, meter, compression hose, bracelet, rollator, cpap |
| Legal | Mlp, legal, court, jail, law, lawyer, crime victims assist, police, prison |

eAppendix D. Other lexical categories used in rule-based algorithm

| Lexical category | Key terms |
|---|--|
| Social Worker terms | social worker, soc worker, soc. worker, s worker, s. worker, sw, social service coordinator, writer, ssc |
| Intervention (verb) terms | provided, informed, suggested, contacted, talked, sent, offered, assessed, discussed, inquired, assisted, suggested, delivered, informed, asked, contacted, updated, collaborated, demonstrated, gave, patient was given, pt was given, pt. was given, reviewed, explained, verified, spoke, notified, consulted, linked, educated, helped, requested, referred, reached, called, placed |
| Forward negative assertion terms for intervention verb terms | social worker, soc worker, soc. worker, s worker, s. worker, sw, social service coordinator, writer, ssc, by patient, by pt, by pt., by mom, by mum, by doctor, by nurse, by rn, by dad, by mother, by father, by son, by daughter, by friend, by uncle, by aunt |
| Backward negative assertion terms for intervention verb terms | social worker was, sw was, social service coordinator was, ssc was, writer was, soc worker was, s. worker was, s worker was, social worker was, patient, pt, pt., mom, doctor, nurse, rn, dad, mother, father, son, daughter, friend, uncle, aunt, not, n't, nt |

eAppendix E. Inter-rater agreement for manual coding

| Interventions | Kappa |
|-----------------------------------|--------------|
| Financial Planning | 0.65 |
| Care coordination | 0.64 |
| Community Service | 0.76 |
| Education | 0.52 |
| Supportive Counseling | 0.66 |
| Filing Applications and Reporting | 0.77 |
| Housing | 0.69 |
| Transportation | 0.69 |
| Durable Medical Equipment | 0.73 |
| Legal | 0.76 |

eAppendix F.

A. 5-fold Cross validation metrics for logistic regression and kernelized SVM algorithms

| Performance Metrics | Logistic Regression | | | | | Kernelized Support Vector Machine (SVM) | | | | |
|-----------------------------------|---------------------|-----------------|----------------------|----------|------|---|-----------------|----------------------|----------|------|
| | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | AUC | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | AUC |
| Financial Planning | 0.96 | 0.98 | 0.82 | 0.89 | 0.98 | 0.97 | 0.98 | 0.88 | 0.93 | 0.97 |
| Care coordination | 0.94 | 0.94 | 0.77 | 0.84 | 0.97 | 0.94 | 0.93 | 0.77 | 0.84 | 0.97 |
| Community Service | 0.94 | 0.93 | 0.69 | 0.79 | 0.94 | 0.93 | 0.93 | 0.68 | 0.78 | 0.94 |
| Education | 0.92 | 0.88 | 0.70 | 0.77 | 0.91 | 0.91 | 0.86 | 0.69 | 0.76 | 0.91 |
| Supportive Counseling | 0.94 | 0.94 | 0.66 | 0.77 | 0.95 | 0.94 | 0.91 | 0.66 | 0.76 | 0.94 |
| Filing Applications and Reporting | 0.95 | 0.89 | 0.61 | 0.72 | 0.94 | 0.94 | 0.86 | 0.61 | 0.71 | 0.94 |
| Housing | 0.98 | 0.97 | 0.73 | 0.82 | 0.99 | 0.98 | 0.93 | 0.73 | 0.81 | 0.98 |
| Transportation | 0.96 | 1.00 | 0.58 | 0.73 | 0.97 | 0.95 | 0.94 | 0.58 | 0.71 | 0.98 |
| Durable Medical Equipment | 0.98 | 0.97 | 0.60 | 0.70 | 0.93 | 0.98 | 0.97 | 0.60 | 0.70 | 0.93 |
| Legal | 0.99 | 1.00 | 0.55 | 0.66 | 0.91 | 0.99 | 1.00 | 0.55 | 0.66 | 0.93 |

B: 5-fold Cross validation metrics for multinomial NB and linear SVM algorithms

| Interventions | Naive Bayes Model | | | | | Linear Support Vector Machine (SVM) | | | | |
|--------------------|-------------------|-----------------|----------------------|----------|------|-------------------------------------|-----------------|----------------------|----------|------|
| | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | AUC | Accuracy | Precision (PPV) | Recall (sensitivity) | F1-score | AUC |
| Financial Planning | 0.85 | 0.58 | 0.75 | 0.65 | 0.86 | 0.89 | 0.73 | 0.68 | 0.74 | 0.85 |
| Care coordination | 0.83 | 0.60 | 0.83 | 0.69 | 0.89 | 0.88 | 0.78 | 0.73 | 0.74 | 0.89 |

| | | | | | | | | | | |
|-----------------------------------|------|------|------|------|------|------|------|------|------|------|
| Community Service | 0.81 | 0.48 | 0.82 | 0.60 | 0.85 | 0.87 | 0.64 | 0.65 | 0.60 | 0.85 |
| Education | 0.79 | 0.51 | 0.69 | 0.58 | 0.80 | 0.82 | 0.66 | 0.68 | 0.62 | 0.82 |
| Supportive Counseling | 0.89 | 0.65 | 0.61 | 0.63 | 0.82 | 0.92 | 0.70 | 0.69 | 0.68 | 0.82 |
| Filing Applications and Reporting | 0.87 | 0.49 | 0.69 | 0.57 | 0.85 | 0.91 | 0.67 | 0.55 | 0.57 | 0.85 |
| Housing | 0.96 | 0.69 | 0.45 | 0.53 | 0.84 | 0.96 | 0.73 | 0.56 | 0.59 | 0.87 |
| Transportation | 0.91 | 0.61 | 0.54 | 0.57 | 0.85 | 0.94 | 0.71 | 0.64 | 0.68 | 0.86 |
| Durable Medical Equipment | 0.96 | 0.51 | 0.34 | 0.40 | 0.87 | 0.98 | 0.71 | 0.63 | 0.60 | 0.87 |
| Legal | 0.97 | 0.13 | 0.10 | 0.11 | 0.63 | 0.97 | 0.21 | 0.23 | 0.26 | 0.66 |

eAppendix G. Patient demographic characteristics by SW intervention type

| Characteristic | All Categories | Financial Planning | Care coordination | Community Service | Educa tion | Supportive Counseling | Applications and Reporting | Housing | Transportation | Durable Medical Equipment | Legal |
|-------------------------------|----------------|--------------------|-------------------|-------------------|-------------|-----------------------|----------------------------|-------------|----------------|---------------------------|-------------|
| Age, mean (SD) | 38.7 (21.4) | 36.6 (18.9) | 35.5 (21.9) | 39.5 (20.8) | 38.9 (20.6) | 36.9 (22.0) | 33.8 (19.1) | 32.9 (18.5) | 34.7 (19.9) | 39.3 (19.7) | 42.9 (17.9) |
| Sex, (%) | | | | | | | | | | | |
| Female | 63.5 | 65.7 | 60.6 | 59.4 | 61 | 57.3 | 60 | 70.6 | 65.9 | 57.1 | 64.7 |
| Male | 36.5 | 34.3 | 39.4 | 40.6 | 39 | 42.7 | 40 | 29.4 | 34.1 | 42.9 | 35.3 |
| Race-Ethnicity, (%) | | | | | | | | | | | |
| Hispanic | 42.9 | 46.3 | 42.6 | 37.7 | 51.2 | 45.3 | 46.7 | 41.2 | 39 | 42.9 | 47.1 |
| African American Non-Hispanic | 35.8 | 31.3 | 36.2 | 39 | 34.2 | 37.3 | 28.9 | 35.3 | 36.6 | 19 | 35.3 |
| White Non-Hispanic | 18.6 | 20.9 | 19.2 | 20.3 | 13.4 | 10.7 | 22.2 | 23.5 | 22 | 28.6 | 17.6 |
| Other Non-Hispanic | 1 | 0 | 1 | 1.5 | 1.2 | 1.7 | 2.2 | 0 | 2.4 | 9.5 | 0 |
| Unknown Race or Ethnicity | 1.7 | 1.5 | 1 | 1.5 | 0 | 4 | 0 | 0 | 0 | 0 | 0 |
| Elixhauser score, mean (S.D) | 3.14 (2.6) | 2.4 (2.3) | 3.4 (2.6) | 3.1 (2.2) | 3.1 (2.5) | 3.2 (2.6) | 2.6 (2.4) | 2.6 (2.4) | 2.8 (2.1) | 3.2 (2.2) | 3.9 (2.8) |
| Total, n | 408 | 67 | 94 | 69 | 82 | 75 | 45 | 17 | 41 | 21 | 17 |

eAppendix H. Average accuracy for all classification algorithms

| Model | Macro-average Accuracy |
|--|-------------------------------|
| Multilabel (On Vs Rest) kernelized SVM with radial basis function | 0.97 |
| Multilabel (One Vs Rest) Logistic Regression | 0.96 |
| Linear SVM | 0.95 |
| Multinomial Naïve Bayes classifier | 0.92 |
| Multilabel LSTM | 0.88 |
| Rule-based | 0.87 |
| Multilabel (Label Powerset) kernelized SVM with radial basis function | 0.82 |
| Multilabel (Classifier Chains) kernelized SVM with radial basis function | 0.81 |
| Multilabel (Binary Relevance) kernelized SVM with radial basis function | 0.8 |
| Multilabel (Binary Relevance) Gaussian Naïve Bayes Classifier | 0.79 |
| Multilabel K-Nearest Neighbor | 0.41 |