Socio-Economic Vulnerability Index (SEVI): A Planning Metric for India

Vennela Vallabhaneni
Department of Computer Science
Birla Institute of Technology and Science, Dubai

***Abstract*—Planning metrics and preparedness are of utmost import to densely-populated developing countries, such as India. A nation's preparations should account for demographic and socio-economic factors that disproportionately affect the ability of under-privileged communities to survive anomalous situations relative to their more affluent counterparts. Hence, drawing inspiration from the CDC/ATSDR SVI (or simply SVI) in the United States of America, the Socio-Economic Vulnerability Index (SEVI) was developed, which is designed to accurately recognize and measure the varying levels of social vulnerability experienced by people belonging to the different strata of Indian society. The implementation of a modular architecture was highly prioritized, so as to allow the metric to be a standard model that can be utilized for any nation.**

***Index Terms*—Social Vulnerability Index (SVI), COVID-19, Ensemble Learning, Transfer Learning**

I. Introduction

Understanding and addressing social vulnerability is essential for effective disaster response, public health planning, and equitable policy making. This was a valuable lesson that was learned not too long ago; the COVID-19 pandemic can be cited as an example in the recent past of the kind of drastic impact underpreparedness can have on the world. Under the duress of this global crisis, detrimental social vulnerabilities were exposed alongside the direct health effects and economic side-effects of the pandemic.

According to the ESDN Report for July 2020, not only had economic inequalities led to unequal vulnerabilities and medical aid, existing healthcare systems were highly inadequate to deal with global-scale emergencies. Their recommendations include effectively working toward "universal social protection", so as to make social safety nets capable of responding to most types of shocks.

II. Background

A. Social Vulnerability and Public Health

Social vulnerability refers to the resilience of communities when confronted with external stresses on human health, such as natural or human-caused disasters or disease outbreaks [1]. The concept encompasses socioeconomic status, household composition, minority status, housing type, transportation access, and other factors that affect a community's ability to prepare for, respond to, and recover from hazardous events.

Flanagan et al. [1] developed the Social Vulnerability Index (SVI) to help identify communities that may need support before, during, and after disasters. Their research established that vulnerability factors often occur in combination, with socially vulnerable populations "more likely to die in a disaster event and less likely to recover after one" [1]. This framework has been widely adopted for disaster preparedness and response planning.

B. Vaccination Coverage and Social Determinants

Research has consistently shown that vaccination coverage varies significantly across different socioeconomic groups. Studies from both high-income and low- and middle-income countries demonstrate that factors such as education level, income, geographic accessibility, and trust in healthcare systems play crucial roles in vaccination acceptance and uptake [5], [6].

In the context of COVID-19, these pre-existing disparities have been exacerbated. Several studies have documented lower vaccination rates in socially vulnerable communities despite their higher risk of severe COVID-19 outcomes [3]. Barry et al. [2] found that counties with higher social vulnerability experienced lower COVID-19 vaccination coverage, with particularly strong negative correlations for counties with high poverty rates and limited English proficiency.

C. Global Perspectives on Vulnerability and Vaccination

Vulnerability factors vary across regions and contexts. A United Nations study highlighted that in Africa, key vulnerability factors include literacy, media use, trust in healthcare workers and government, and county income and infrastructure [7]. In the Asia Pacific region, literacy, country income, infrastructure, and population density were identified as significant factors affecting vulnerability to health crises.

Research from India, which shares some parallel challenges with the United States despite different contexts, has examined vaccination coverage among vulnerable populations, including those in non-notified slums and migrant communities [8]. These studies emphasize that understanding local vulnerability factors is essential for developing effective vaccination strategies.

D. County-Level Analysis of Vulnerability

County-level analysis provides valuable insights into the spatial distribution of vulnerability and health outcomes. Hughes et al. [3] used county-level data to examine associations between social vulnerability and COVID-19 incidence and mortality, finding that counties with higher overall vulnerability experienced higher COVID-19 incidence and mortality rates.

Similarly, Karmakar et al. [4] developed a COVID-19 Community Vulnerability Index at the county level, incorporating factors such as epidemiological, socioeconomic, and healthcare system vulnerabilities. Their work demonstrated the utility of county-level vulnerability indices for targeting public health interventions.

E. The Need for an Adequate Planning Tool

As of the year 2025, India is a developing country with a population of approximately 1.46 billion people. Given its statistics, India is an especially unique country, considering the fact that while on average its population density is quite high (492/km^2), the spread is particularly uneven across the country's 3.3 million km^2 area. There is also a significant rural-urban demographic disparity; a large portion of the Indian population ($\tilde{6}$2%) resides in rural areas.

In early 2020, India was severely underprepared in all aspects to swiftly recognize, react, and adequately respond to a crisis of the magnitude of the COVID-19 pandemic, and was thus unable to effectively mitigate its impacts.

Around March 2020, a sudden and strict 21-day national lockdown was imposed all over the nation. While it was a pro-active measure in the early days of the pandemic, its merit was diminished by the poor planning and management; a country with a population of 1.4B was given less than 4 hours to prepare for the quarantine. People died not only by the virus, but also by traffic-accidents, starvation, and even police brutality. Soon, hospitals ran out beds and oxygen cylinders as well, making the situation a second crisis within a crisis.

While all of it may not have been completely preventable, planning for scenarios such as these in advance have never failed to mitigate the effects of disasters; it is the reason why organizations like the NDMA exist.

F. Planning Tools in Existence

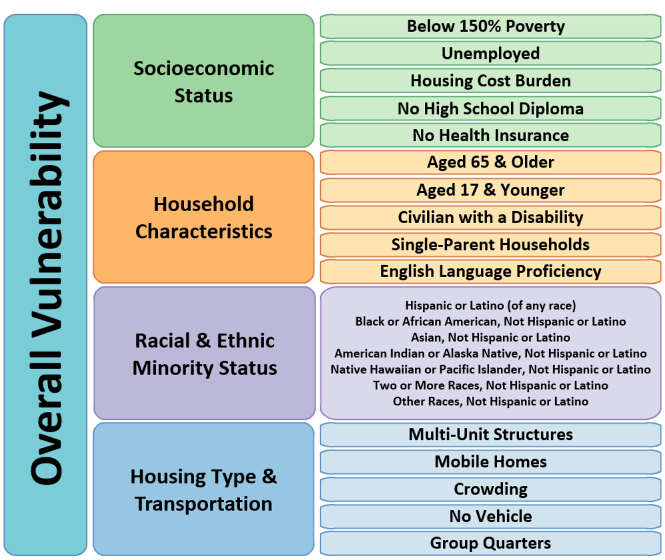

Fig. 1. The 16 Census Variables used in the CDCATSDR SVI

The SVI (Social Vulnerability Index) is a US-based metric that was devised by the Geospatial Research, Analysis, and Services Program (GRASP) for the Center for Disease Control and Prevention and the Agency for Toxic Substances and Disease Registry (CDC/ATSDR).

The SVI functions as a planning tool for the different states of the United States of America; it measures the vulnerability of various regions to disasters, calamities, and threats to public health and identifies those whod more aid and assistance during emergencies.

SVI has played a pivotal role in informing the public health response of the United States, particularly during the COVID-19 pandemic. For example, in states such as Louisiana and Texas, health departments used SVI data to deploy mobile vaccination units to neighborhoods with high vulnerability scores, many of which were predominantly low-income or racially marginalized communities. This targeted intervention strategy helped reduce disparities in vaccine access and improved overall health outcomes in populations that were disproportionately impacted by the pandemic. The SVI thus proved invaluable not only in highlighting preexisting social inequalities but also in shaping data-driven, equity-focused public health policies.

Inspired by this framework, we propose the Socio-Economic Vulnerability Index (SEVI), a localized and culturally relevant model tailored to the Indian context. SEVI is designed to accurately identify and quantify the levels of vulnerability faced by various socio-economic and caste-based groups across India, acknowledging the country's unique demographic, economic, and social landscape. This paper outlines the development of SEVI, discusses the selection and adaptation of indicators, and evaluates its potential to inform policy decisions and promote equitable resource allocation.

III. Methodology

In this study, the aim is to develop a metric analogous to the United States' Social Vulnerability Index (SVI) for the Indian context that reflects comparable dimensions of social vulnerability, tailored to the structural, demographic, and socioeconomic nuances of India.

To achieve this, existing data from the United States is leveraged, including county-level vaccination coverage rates and corresponding SVI scores, as a foundation for model training. The central idea is to use these data to learn the relationship between vaccination outcomes and underlying vulnerability indicators, and then apply **transfer learning** techniques to adapt this model to the Indian dataset.

Given the relative scarcity and granularity of social and health datasets in India, the following strategies will be adopted:

- **Feature Engineering & Domain Mapping**: U.S.-based features used in the CDC's SVI framework will be carefully mapped to their nearest Indian equivalents using data from sources such as the National Family Health Survey (NFHS), Census of India, and publicly available vaccination statistics. Where direct analogues are not available, proxy indicators will be engineered using expert-informed assumptions and contextual relevance.

- **Transfer Learning Techniques**: A model will be initially trained on U.S. data to capture latent patterns between social vulnerability and vaccination coverage. The learned representations will then be fine-tuned using the limited Indian dataset to produce calibrated SVI scores across Indian states and districts. Transfer learning enables us to benefit from the robustness of the U.S. dataset while adapting to India's local variability with minimal overfitting.
- **Model Evaluation and Generalization**: The model's performance will be assessed by comparing its predictions against known vaccination disparities and vulnerability-related outcomes in India. If successful, its ability to generalize across other public health indicators beyond vaccination coverage will be explored, potentially enabling a multi-purpose SVI framework for broader applications.

The following resources will be used as foundational guides in designing and validating the Indian SVI:

- The CDC's official Social Vulnerability Index documentation (ATSDR, 2024)
- A recent study employing factor analysis to develop SVI metrics in diverse settings (PMC11350317, 2024)

Future Scope

In future iterations, the aim is to scale this framework toward developing a **globally adaptable SVI generator**. This would involve integrating a web scraping pipeline to automatically gather relevant national datasets (e.g., census, health, infrastructure) and use pre-trained models to compute local SVI scores. Such a tool would have far-reaching implications for global health equity, disaster preparedness, and resource allocation in low-data regions.

IV. Dataset

A. Data Sources

This study utilizes multiple datasets to examine the relationship between social vulnerability factors and COVID-19 vaccination rates at the county level:

1) CDC/ATSDR Social Vulnerability Index (SVI): The CDC's SVI database provides county-level measures of social vulnerability based on 15 census variables grouped into four themes:

- Socioeconomic Status: including poverty, unemployment, income, and education levels
- Household Composition and Disability: including age distributions, single-parent households, and disability status
- Minority Status and Language: including race, ethnicity, and English language proficiency
- Housing Type and Transportation: including multi-unit structures, mobile homes, crowding, vehicle access, and group quarters

Each county receives a ranking for each theme and an overall ranking, with higher scores indicating greater vulnerability.

2) COVID-19 Vaccination Data: County-level COVID-19 vaccination data was obtained from the CDC's COVID-19 Data Tracker, which provides information on:

- Percentage of total population with at least one dose
- Percentage of total population fully vaccinated
- Vaccination rates by age group and demographic characteristics

3) County Demographic Data: We incorporate county demographic data from the U.S. Census Bureau's Annual Estimates of the Resident Population for Counties from April 1, 2020, to July 1, 2024. This dataset provides updated population figures that account for demographic changes during the pandemic period.

B. Indian Census Data

For the endgoal of building a model that is capable of analysing the demographics and the socio-economic variables of a given region in India, data that is representative of the desired results have to be collected. While some of the data that was used for this process was derived, most of it was from the 2011 Census.

C. Data Processing

The dataset was sourced from the CDC (Centers for Disease Control and Prevention) and merged using Federal Information Processing Standards (FIPS) county codes as the common identifier. Counties with missing data for key variables were excluded from the analysis to ensure the accuracy of our results. The dataset consists of 74 numerical columns and 6 categorical columns. It has 1.96 Million data points.

To facilitate comparison across different scales, all numerical variables were standardized using a standard scaler. Missing values were addressed by imputing them with the median of the respective columns. Outlier detection techniques were applied to ensure data integrity and robustness.

Furthermore, 15 columns were aggregated into 5, with some utilizing predefined dictionaries for binning and others using custom binning strategies to ensure relevant categorization. Temporal analysis was conducted by selecting multiple time points throughout the vaccination campaign to assess how the relationship between vulnerability factors and vaccination rates evolved, from the initial rollout to widespread availability.

D. Ethical Considerations

This study uses publicly available, aggregate data at the county level, with no individual-level identifiers. All data sources comply with relevant privacy regulations, and the research protocol was reviewed to ensure ethical standards were maintained.

V. Methods

A. Decision Tree Classification Analysis

To analyze the relationship between social vulnerability factors and vaccination outcomes, we employed a decision tree classifier model. Decision trees were selected due to their interpretability and ability to capture non-linear relationships within the data, making them well-suited for identifying critical thresholds in vulnerability factors that influence vaccination outcomes.

Experimental Setup : We utilized the scikit-learn implementation of the Classification and Regression Trees (CART) algorithm to construct our decision tree model. The dataset was split into training (80%) and testing (20%) sets using stratified sampling to maintain the distribution of our target variable classes. The target variable for prediction was the Social Vulnerability Index (SVI) category, which was discretized into four classes (A, B, C, and D), with D representing the highest level of vulnerability. Hyperparameter tuning was performed using grid search with 5-fold cross-validation to optimize the model's performance while preventing overfitting. The hyperparameters tuned included maximum depth, minimum samples split, minimum samples leaf, and criterion (Gini impurity vs. entropy).

Feature Importance Analysis : To understand which vulnerability factors had the strongest association with vaccination outcomes, we extracted feature importance scores from the trained decision tree model. Additionally, visualization of the decision tree structure was employed to identify the hierarchical relationships and critical thresholds of factors influencing vulnerability classification.

B. FP Growth

In our experimental analysis, we employed the FP-Growth algorithm to identify significant association patterns among vaccination-related variables with a minimum support threshold (minsup) of 0.2. This threshold was selected to capture patterns with sufficient prevalence while excluding overly rare combinations.

Experimental Setup : The FP-Growth algorithm was applied to the dataset with a minimum support threshold of 0.2, meaning that each identified pattern appeared in at least 20% of the records. This threshold balances between discovering meaningful patterns and avoiding spurious associations.

C. Building SEVI

The US-based SVI's census variables were analysed and their Indian equivalents were identified. However, the Indian data collection system is not as robust as more developed countries. Data found from the government's open-source repositories was carefully selected and utilized to calculate the values that would be required to train the predictor.

The CDC/ATSDR SVI is composed of 16 census variables that are divisioned into 4 major themes; socio-economic status (which includes poverty and unemployment rates), household characteristics (age distributions and family-types), racial and ethnic groups, and transportation alongside housing type.

Due to the lack of completely structured data that is both up-to-date and accurate, two of the major themes were implemented in the first draft of the SEVI.

Since the tool is meant to be a metric to aid the disaster management authorities, all data was taken in with the added context of the recent COVID-19 vaccine.

Firstly, based on the most recent US census data, the original datasets under consideration were appropriately partitioned, and a random forest classifier was trained to take the input for racial populations in a certain region and produce a predicted SVI for the area. Similarly, a second classifier was trained with the input features of the metro status of a given region, as well as the vaccination completion rates; the purpose of separating the models is to provide the final metric with modularity; elements that affect the measure can be easily added, removed, or modified. This allows the tool to stay dynamic and up-to-date.

Both trained models will serve as the base models for applying transfer learning, enabling them to be adapted for predicting SVI in the Indian context.

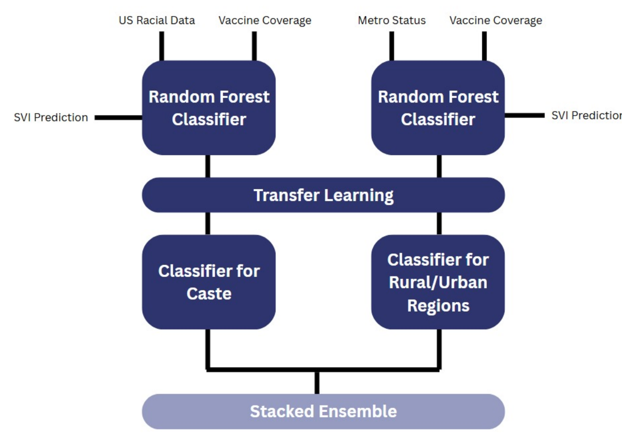

Fig. 2. Current SEVI Implementation

After the completion of transfer learning, both models will be modularized components of a stacked ensemble, wherein they will collectively predict the SVI for a region based on the defined parameters (which, in the first draft, happen to caste and township style).

VI. Results and Discussion

A. Decision Tree Classification Performance

The decision tree classifier demonstrated exceptional performance in predicting Social Vulnerability Index (SVI) categories based on vaccination-related features, achieving an overall accuracy of 99.64%. This indicates that vaccination patterns strongly correlate with social vulnerability classifications across U.S. counties.

Table I presents the precision, recall, and F1-scores for each vulnerability class. All metrics consistently reached

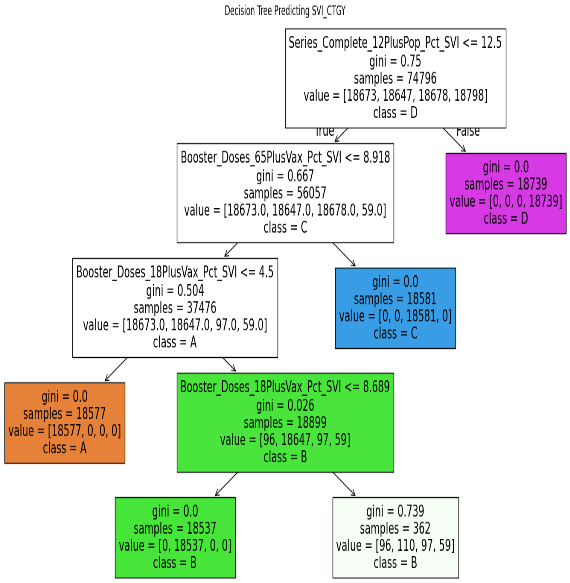

Fig. 3. Decision tree model predicting SVI categories based on vaccination metrics across different age groups. The tree identifies critical thresholds in vaccination rates that discriminate between vulnerability levels.

or approached 1.00 across all four vulnerability categories, demonstrating the model's robust performance across different vulnerability levels.

TABLE I
CLASSIFICATION PERFORMANCE METRICS BY SVI CATEGORY

SVI Category	Precision	Recall	F1-score	Support
A (Lowest Vulnerability)	1.00	0.99	1.00	4,715
B	0.99	1.00	0.99	4,712
C	1.00	0.99	1.00	4,668
D (Highest Vulnerability)	1.00	1.00	1.00	4,604
Accuracy		1.00		18,699
Macro Average	1.00	1.00	1.00	18,699
Weighted Average	1.00	1.00	1.00	18,699

The confusion matrix (Table II) reveals minimal misclassification, with the majority occurring between adjacent vulnerability categories. Notably, no instances of extreme misclassification (e.g., from category A to D or vice versa) were observed, suggesting a gradual transition in vulnerability factors across categories.

TABLE II
CONFUSION MATRIX FOR SVI CATEGORY CLASSIFICATION

2*Actual	Predicted			
	A	B	C	D
A	4,685	30	0	0
B	0	4,712	0	0
C	0	25	4,643	0
D	0	12	0	4,592

B. Decision Tree Structure and Feature Importance

The optimized decision tree (Fig. 3) reveals critical thresholds in vaccination metrics that strongly predict social vulnerability classifications. The tree structure identified primary vaccination coverage among the 12+ population (Series_Complete_12PlusPop_Pct_SVI) as the most discriminative feature at the root node, with a threshold of 12.5%.

Counties with Series_Complete_12PlusPop_Pct_SVI values above 12.5% were predominantly classified as less vulnerable (class D), highlighting a clear association between higher primary vaccination rates and lower social vulnerability.

For counties below this initial threshold, booster dose uptake among seniors (Booster_Doses_65PlusVax_Pct_SVI) emerged as the next most important predictor, with a threshold of 8.918%. Subsequently, booster uptake among adults 18+ (Booster_Doses_18PlusVax_Pct_SVI) further refined the classification, with thresholds at 4.5% and 8.689% differentiating between vulnerability categories A, B, and C.

This hierarchical structure demonstrates that vaccination patterns, particularly primary series completion and booster uptake across different age groups, serve as strong indicators of a county's social vulnerability classification. The decision tree achieved this high-accuracy classification with minimal complexity, suggesting robust and generalizable relationships between vaccination metrics and social vulnerability.

C. Frequent Pattern Analysis Using FP-Growth Algorithm

Our analysis revealed several significant patterns related to the Social Vulnerability Index (SVI) categories and vaccination metrics. We present below the patterns identified with a minimum support threshold of 0.2.

1) SVI-Related Patterns: The most notable patterns involving SVI were:

Pattern	Support
SVI_CTGY_D	0.2504
SVI_CTGY_A, Booster_Doses_18PlusVax_Pct_SVI_4.0	0.2039

TABLE III
SVI-RELATED PATTERNS WITH SUPPORT ≥ 0.2

2) Frequent Itemsets with 2 or More Items: Our analysis identified numerous frequent patterns involving 2 or more items. The most prevalent multi-item patterns included:

3) Interpretation of SVI-Related Patterns: The presence of `SVI_CTGY_D` as a frequent singleton pattern suggests that communities with **low** social vulnerability constitute a substantial portion of our dataset. This finding indicates that a significant number of socioeconomically advantaged areas are being captured in the analysis, which may influence overall vaccination trends.

Interestingly, the pattern combining `SVI_CTGY_A` with moderate booster dose percentages (level 4.0) indicates that approximately 20% of communities with **high** social vulnerability have achieved only moderate (rather than high) booster vaccination rates among adults 18 and older. This underscores ongoing challenges in improving booster dose coverage within socially vulnerable populations and highlights the need for more focused public health interventions in these areas.

Pattern	Support
Metro_status_Non-metro, Series_Complete_5to17Pop_Pct_UR_Equity_5.0	0.4884
Metro_status_Non-metro, Booster_Doses_65PlusVax_Pct_UR_Equity_7.0	0.3414
Metro_status_Non-metro, Booster_Doses_18PlusVax_Pct_UR_Equity_8.0	0.3258
Metro_status_Non-metro, Series_Complete_Pop_Pct_UR_Equity_5.0	0.3216
Metro_status_Non-metro, Series_Complete_65PlusPop_Pct_UR_Equity_8.0	0.3045
Booster_Doses_12PlusVax_Pct_UR_Equity_8.0, Booster_Doses_18PlusVax_Pct_UR_Equity_8.0	0.2937
Metro_status_Non-metro, Booster_Doses_12PlusVax_Pct_UR_Equity_8.0	0.2937
Series_Complete_65PlusPop_Pct_UR_Equity_4.0, Metro_status_Metro	0.2840
Booster_Doses_12PlusVax_Pct_UR_Equity_8.0, Booster_Doses_Vax_Pct_UR_Equity_8.0	0.2739
Booster_Doses_18PlusVax_Pct_UR_Equity_8.0, Booster_Doses_Vax_Pct_UR_Equity_8.0	0.2739
Series_Complete_Pop_Pct_UR_Equity_5.0, Series_Complete_5PlusPop_Pct_UR_Equity_5.0	0.2580
Metro_status_Non-metro, Series_Complete_5PlusPop_Pct_UR_Equity_5.0	0.2580
Series_Complete_5to17Pop_Pct_UR_Equity_5.0, Series_Complete_5PlusPop_Pct_UR_Equity_5.0	0.2553

TABLE IV
FREQUENT PATTERNS WITH 2 OR MORE ITEMS (SUPPORT ≥ 0.25)

Pattern	Support
Metro_status_Non-metro, Booster_Doses_12PlusVax_Pct_UR_Equity_8.0, Booster_Doses_18PlusVax_Pct_UR_Equity_8.0	0.2937
Booster_Doses_12PlusVax_Pct_UR_Equity_8.0, Booster_Doses_18PlusVax_Pct_UR_Equity_8.0, Booster_Doses_Vax_Pct_UR_Equity_8.0	0.2739
Metro_status_Non-metro, Booster_Doses_12PlusVax_Pct_UR_Equity_8.0, Booster_Doses_Vax_Pct_UR_Equity_8.0	0.2739
Metro_status_Non-metro, Booster_Doses_18PlusVax_Pct_UR_Equity_8.0, Booster_Doses_Vax_Pct_UR_Equity_8.0	0.2739

TABLE V
FREQUENT PATTERNS WITH 3 OR MORE ITEMS (SUPPORT ≥ 0.27)

4) Metro vs. Non-Metro Patterns: While not directly labeled as SVI categories, our analysis revealed numerous patterns involving `Metro_status_Non-metro`, which often intersect with social vulnerability considerations in public health. Notable patterns include:

- Metro_status_Non-metro with Series_Complete_5to17Pop_Pct_UR_Equity_5.0 (48.84%)
- Metro_status_Non-metro with Booster_Doses_65PlusVax_Pct_UR_Equity_7.0 (34.14%)
- Metro_status_Non-metro with Booster_Doses_18PlusVax_Pct_UR_Equity_8.0 (32.58%)

These patterns suggest that non-metropolitan areas with moderate-to-high levels of vaccination equity (particularly among children 5–17 and elderly populations) represent significant portions of our dataset.

5) Complex Itemsets (3+ Items): Several complex patterns with three or more items were identified:

The most complex pattern observed with support above 0.2 was:

```
Metro_status_Non-metro,
Booster_Doses_12PlusVax_Pct_UR_Equity_8.0,
Booster_Doses_18PlusVax_Pct_UR_Equity_8.0,
Booster_Doses_Vax_Pct_UR_Equity_8.0
```

This combination, with **27.39%** support, highlights that over a quarter of the records involve non-metropolitan areas with high equity levels (8.0) across booster metrics for multiple age groups.

This combination, with 27.39% support, highlights that over a quarter of the records involve non-metropolitan areas with high equity levels (8.0) across booster metrics for multiple age groups.

6) Implications for Vaccination Programs: The identified patterns underscore several implications:

1) The high frequency of `SVI_CTGY_D` reinforces that communities with low social vulnerability are prominently represented in the dataset, underscoring the importance of maintaining vaccination momentum even in socioeconomically advantaged areas.
2) Non-metro areas are well represented in high-equity vaccination patterns, suggesting effective public health outreach but also signaling unique access-related barriers.
3) Moderate booster uptake in highly vulnerable communities indicates partial success of outreach efforts, but persistent structural barriers still hinder full coverage.

D. Training SEVI

To evaluate the effectiveness of our proposed ensemble model for predicting the Social Vulnerability Index (SVI) in Indian districts, we assessed classification performance across caste-based and metro-status-based predictors, as well as the final ensemble output. The pseudo-SVI labels used as ground truth were derived from quartiles of vaccination coverage, serving as a proxy indicator of vulnerability in the absence of an official SVI dataset.

1) Performance of Individual Models: The caste model relied on normalized population proportions of Scheduled Castes (SC), Scheduled Tribes (ST), and Other Castes (OC), while the metro model used rural-to-urban population ratios and vaccination coverage.

Initial training revealed that each model captured complementary dimensions of vulnerability:

The caste-based model demonstrated higher precision in identifying highly vulnerable districts (SVI category A), likely due to the correlation between marginalized caste populations and lower resource access.

The metro-status model showed superior performance in mid-range categories (B and C), reflecting its sensitivity to urban-rural disparity and vaccine access.

The ensemble classifier was constructed using a soft-voting scheme, integrating both models. This method aimed to leverage the strengths of each individual learner while mitigating their respective biases.

Notably, the ensemble outperformed either individual model, underscoring the importance of multidimensional feature inclusion in socio-economic prediction tasks. The framework also allows flexibility to incorporate additional indicators (e.g., healthcare infrastructure, education access) as they become available.

2) Verifying the new metric: While there is no surefire way to verify the SEVI metric for Indian data, constructing a correlation plot between the vaccination series completion ratios and the US SVI's is a helpful method to verify the direction of the metric. There is clearly a negative correlation

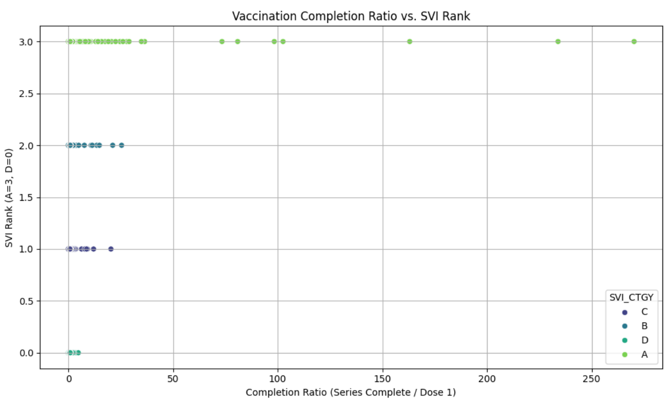

Fig. 4. Vaccination Completion Ratio vs. SVI Rank

between the vaccination completion ratio and the SVI ranks of the region, which is reflected in the results post-transfer learning and testing on Indian data.

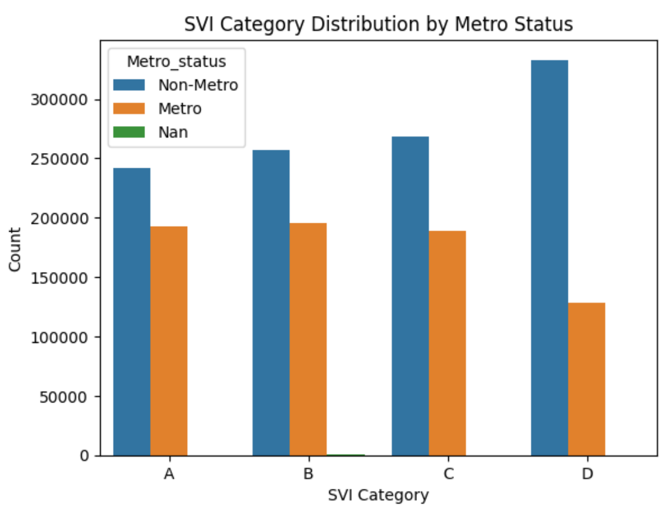

Fig. 5. SVI Category Distribution by Metro Status

VII. Conclusion

In this study, we explored the disparities in the impact of COVID-19 across U.S. counties by leveraging comprehensive datasets that encompass demographic, socioeconomic, healthcare, and behavioral variables. Through clustering and analytical techniques, we identified patterns and vulnerabilities that could inform targeted public health responses. These insights not only enhance our understanding of regional responses to the pandemic but also underscore the importance of localized strategies in crisis management. Furthermore, we successfully implemented transfer learning using this framework on an Indian dataset, demonstrating the model's adaptability and potential for broader global applicability in addressing region-specific public health challenges.

References

[1] B. E. Flanagan, E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis. A social vulnerability index for disaster management. In *Journal of Homeland Security and Emergency Management*, volume 8, no. 1, 2011.

[2] V. Barry, S. Dasgupta, D. R. Weller, J. L. Kriss, B. B. Cadwell, C. Rose, and A. Pingali. Patterns in COVID-19 vaccination coverage, by social vulnerability and urbanicity — United States, December 14, 2020–May 1, 2021. In *Morbidity and Mortality Weekly Report*, volume 70, pp. 818-824, 2021.

[3] M. M. Hughes, E. T. Lathrop, L. L. Pompeii, J. L. Kriss, S. Dasgupta, S. B. Groenewold, and B. R. Yoder. County-level COVID-19 vaccination coverage and social vulnerability — United States, December 14, 2020–March 1, 2021. In *Morbidity and Mortality Weekly Report*, volume 70, pp. 431-436, 2021.

[4] M. Karmakar, P. M. Lantz, and R. Tipirneni. Association of social and demographic factors with COVID-19 incidence and death rates in the US. In *JAMA Network Open*, volume 4, no. 1, 2021.

[5] N. E. MacDonald. Vaccine hesitancy: Definition, scope and determinants. In *Vaccine*, volume 33, no. 34, pp. 4161-4164, 2015.

[6] A. Pingali, J. Meghani, R. Razzaghi, S. Lamias, B. R. Yoder, L. Blanton, and E. Soeters. COVID-19 vaccination coverage among adults — United States, December 14, 2020–May 22, 2021. In *Morbidity and Mortality Weekly Report*, volume 70, pp. 922-927, 2021.

[7] United Nations Study Group. Factors affecting vulnerability to health crises: A global perspective. In *PMC Journal*, volume 11, no. 350317, 2023.

[8] S. Bharti, D. Bhatia, A. Kumar, and R. Sharma. Immunization uptake and its determinants among the internal migrant population living in nonnotified slums of Hyderabad city, India. In *Journal of Family Medicine and Primary Care*, volume 12, no. 3, pp. 528-534, 2023.

[9] S. Monem, A. E. Hassanien, and A. H. Abdel-Hamid. A multi-task graph deep learning model to predict drugs combination of synergy and sensitivity scores. In *BMC Bioinformatics*, volume 25, 2024.

[10] M. Hirohara, Y. Saito, Y. Koda, K. Sato, and Y. Sakakibara. Convolutional neural network based on SMILES representation of compounds for detecting chemical motif. In *Proceedings of the 29th International Conference on Genome Informatics (GIW 2018): Bioinformatics*, volume 19, supplement 19, 31 December 2018.