Wildfire Severity and Structural Damage Prediction in Wildland-Urban Interface (WUI) Areas

Chanyoung Park
chp025@ucsd.edu

Reema Alsaeed
ralsaeed@ucsd.edu

Rushyendra Katabathuni
rkatabathuni@ucsd.edu

Jahnavi Naik
janaik@ucsd.edu

Ilyes Meftah
imeftah@cotality.com

Jonathan Sam
jonsam@cotality.com

Abstract

In recent years, California has experienced unprecedented levels of wildfire frequency and intensity. Climate change and various environmental factors, such as droughts and high winds, have been linked to the severe burning of vegetation and destruction of property in regions of high risk. These fires not only affect natural landscapes but are increasingly affecting urban communities, as demonstrated by recent wildfires in California, such as the Palisades fires. Wildfires often leave residents, community planners, and insurance companies struggling to understand why certain structures were destroyed while others survived. This brings us to the focus of this paper, where we construct a model to predict the structural damage on residential structures in Wildland-Urban Interface (WUI) areas using readily available geospatial, fuel, topographic, weather, and structural data. Despite the growing concerns and effects of wildfires on structures, publicly accessible tools that assess damage at the structure level are limited. This lack of structure-level insights makes it more difficult for residents to be prepared, for community planners to identify vulnerable neighborhoods, and for insurance companies to assess risk accurately. To address this gap, we aim to better understand structural damage in WUI areas by using readily available data to train random forest and neural network models and evaluate their performance against ground truth wildfire structure damage data.

Website: https://reemas03.github.io/structural_damage_wui_website/
Code: https://github.com/ChanyoungPark07/structural_damage_wui

1 Introduction

Wildfires have become an increasingly important issue in California, as recent fires have produced high-intensity burns and caused significant threats to nearby communities. Wildfires often leave residents, community planners, and insurance companies struggling to understand why certain structures were destroyed while others survived, which creates uncertainty about the vulnerability of their properties in Wildland-Urban Interface (WUI) areas. Publicly available tools that can provide structural-level insights into wildfire damage are very limited.

In this project, we aim to fill this gap by better understanding structural damage outcomes in WUI areas. With this focus, we can contribute insights that can support better preparedness, safer community planning, and more informed decision-making by residents, planners, and insurers in WUI areas. More specifically, we investigate whether residential structure damage in California wildfires can be predicted using readily available geospatial, fuel, topographic, weather, and structural data with Neural Network and Random Forest models. Our results indicate that both structural and environmental variables contributed meaningfully to predicting residential structure damage. Some of the most influential features included combustible exterior siding, burn severity class, and precipitation. The Random Forest model achieved a macro F1-score of 0.84 and the Neural Network model achieved a macro F1-score of 0.78. These findings suggest that structure-level wildfire damage can be reasonably predicted using publicly available data, although some outcomes remain difficult to capture due to the complexity and variability of wildfire behavior.

Prior research has examined wildfire risk and behavior in California, emphasizing the role of environmental, spatial, and structural variables in shaping wildfire impacts within WUI areas. For example, in ”Assessing wildland-urban interface fire risk,” Mahmoud and Chulahwat (2020) examined community-level wildfire risk in WUI areas. Their study evaluated four different communities and highlighted that wildfire risk is closely related to wind speed and direction, surrounding vegetation patterns, and building layout. By analyzing these factors at a community scale, their work emphasizes the importance of considering both environmental and structural variables in wildfire risk assessment. However, their approach operates at a community scale rather than predicting damage outcomes for individual residential structures. As a result, while their work provides valuable insights into WUI risk patterns, the question of structural-level damage prediction remains less explored.

Similarly, in ”Where wildfires destroy buildings in the US relative to the wildland–urban interface and national fire outreach programs,” Kramer et al. (2018) analyzed buildings threatened and destroyed by wildfire across the United States. Their findings show that the majority of destroyed structures were located within WUI areas. Their work highlights wildfire destruction patterns at a national scale and the increased vulnerability of structures in WUI areas. However, their study focused on identifying spatial patterns of building damage rather than developing predictive models for individual structure damage outcomes.

These studies, along with others done previously, demonstrate that environmental and structural variables contribute significantly to wildfire impacts, yet predictive modeling at a structure-level remains relatively underexplored. Motivated by this gap, our work applies machine learning models to predict residential structure damage in WUI areas using publicly available data.

The datasets in our study provide high-resolution spatial, environmental, and structural information necessary to predict residential structure damage caused by wildfires in California. The Cal Fire Damage Inspection (DINS) reports CAL FIRE (2022) document structures inside or within 100 meters of the fire perimeters. These records include the structure type, construction materials of different structural components, and the ground truth of damage outcomes. In addition, the National Structure Inventory (NSI) dataset Inventory (n.d.) provides supplementary information on individual structures, including structural attributes that complement the DINS dataset. Together, these datasets allow us to incorporate structure-level features with environmental features, which enables us to analyze which types of structures are most vulnerable in WUI areas and which factors are most predictive of structure damage outcomes.

In addition to structure-level inspection data, spatial and environmental data are incorporated through multiple datasets. Wildland-Urban Interface (WUI) boundary data from SILVIS maps SILVIS Lab (n.d.) provide us with spatial information, allowing us to identify the zones where structures interface with wildland vegetation. This ensures our study focuses on structures that fall within WUI boundaries. From our Quarter 1 Project, we have historical fire perimeters and burn severity data from MTBS MTBS (2025) which provide additional wildfire burn severity context. Environmental data from LANDFIRE LANDFIRE (2025), including data such as fuel models and vegetation, help capture conditions that influence wildfire behavior. Land cover information from NLCD Multi-Resolution Land Characteristics Consortium (2025) and weather data from PRISM PRISM Climate Group, Oregon State University (2025) further help to understand the landscape and weather conditions around the fires. By integrating these diverse data sources, we construct a structure-level dataset that contains comprehensive environmental, topographic, meteorological, and structural variables, enabling the development of predictive models for residential structure damage caused by wildfires in WUI areas.

2 Methods

2.1 Data Collection

In this project, we integrated data from five different sources to predict structural damage outcomes in California Wildland-Urban Interface (WUI) areas. The target variable came from the CalFire Damage Inspection (DINS) Dataset, which contained post-fire structural damage assessments across California wildfires. We decided to predict a binary outcome, so we retained only structures that were classified as “Destroyed (>50%)” or “No Damage” in order to focus on the most clear damage states and simplify our predictions.

Our feature data came from four different sources. The first was the merged Fire Severity Data from our Quarter 1 Project, containing information about vegetation, topography, weather conditions, and burn severity at a pixel level. The second was structure-level features from the DINS dataset, such as building type and construction materials. The third was the SILVIS WUI dataset, which provided housing density and vegetation data in WUI areas. Finally, we incorporated US Army National Structure Inventory (NSI) data, which provided building details such as square footage, number of stories, and foundation characteristics.

2.2 Data Preprocessing

DINS Dataset Preparation

We preprocessed the DINS dataset by first standardizing and cleaning column names, removing special characters, converting to lowercase, and replacing spaces with underscores. We then filtered the dataset to only include data about the residential structure types and the structural damage classification of the 30 California wildfires that we selected and decided to explore further. We then reprojected the coordinates from EPSG:6414 to EPSG:5070 to ensure consistency with all of our data. We finally standardized invalid entries to missing and dropped all columns with more than 70% missing values.

Spatial Data Integration

To aggregate the data, multiple spatial joining operations were utilized. First, we clipped the SILVIS WUI data to our fire perimeters to filter out WUI blocks that aren't within our chosen wildfire perimeters. This ensured that our analysis only focused on structures in WUI areas. Next, we merged the cleaned DINS data and the filtered WUI blocks to identify structures that are located within the WUI areas. We standardized fire incident names to ensure proper merging and only retained rows where the DINS incident name matches the Wildfire Perimeter Incident Name. Finally, we performed a spatial join to merge the DINS structures in WUI areas with the Merged Fire Severity Data from Quarter 1. For each structure, we created a 30-meter buffer and found the structures that intersect with the 30-meter pixel-level Fire Severity data.

Next, we rounded out the dataset with NSI building features using a nearest-neighbor spatial join with a maximum distance of 1500 meters. To remove structures with multiple NSI matches, we then retained the closest structural match. Additionally, we engineered a feature, nearest_struct_dist, representing the Euclidean distance between a structure and its nearest neighbor. This feature aimed to provide the models with context on structure density, to help with model predictions. Finally, we removed duplicates based on geometry to ensure each structure only appeared once in our data.

Final Cleaning and Missingness Imputation

To prepare the dataset for our models, final cleaning steps were taken. We removed columns with all false values (from one-hot encoding) and dropped redundant columns such as perimeter metadata and fire incident IDs. By merging in NSI data, we were able to drop two DINS columns (year_built and assessed_improved_value) that had high percentages of missing values, as the NSI data provided more complete and relevant versions of these attributes (med_yr_blt and val_struct, respectively). Categorical variables that represented structural features, including roof construction, exterior siding, vent screens, etc. had missing values imputed with an "Unknown" category. We then one-hot encoded all categorical features, such as structure type, building type, and foundation type. Some columns from the Fire Severity data that were previously one-hot encoded, such as fuel models and land cover, had missing values imputed with False, as the missing values resulted from the merge operations: when a classification category that was present in the Q2 structure data but missing in the Q1 Fire Severity data, the merge would produce missing values for that category because it hadn't seen that column previously. Finally, the target variable was properly converted into a binary variable, with "No Damage" encoded as 0 and "Destroyed (>50%)" encoded as 1. After all preprocessing steps were completed, the final dataset contained 41,769 structure records across 322 features, with the large feature space being caused primarily by one-hot encoding of categorical variables.

To evaluate whether fire characteristics influenced model performance, we trained six separate Random Forest and Neural Network models on different subsets of the data: all fires, small fires (fires with fewer structure records), wind-driven fires, plume driven fires, low-severity fires (fires where less than 33% of pixels had a burn severity of 3 or 4 on the MTBS scale), and high-severity fires (fires where more than 33% of pixels had a burn severity of 3 or 4 on the MTBS scale). Each subset followed the same training, tuning, and threshold optimization as above, including feature selection for each subset.

To prevent overfitting on the dataset and increase the interpretability of the model, multicollinear features were removed, and feature selection was run using the top 95% cumulative importance on a LightGBM for each model training set.

2.3 Random Forest

Model development started with a baseline Random Forest model to classify if structural damage would occur (structure damage over 50%). We then conducted hyperparameter tuning for parameters including `n_estimators`, `max_depth`, `min_samples_split`, `min_samples_leaf`, and `max_features`. To achieve this, we ran `RandomizedSearchCV` with 20 sampled configurations and 3-fold `GroupKFold` cross-validation, grouping by wildfires to maintain event-level separation during training to prevent leakage, and optimizing F-1 score. After selecting the best performing model for each wildfire subset, we performed threshold tuning on the validation set to find the threshold that would maximize the validation F1-score, and evaluated the final model on the test set.

2.4 Neural Network

A feedforward Neural Network was trained on the merged dataset to predict the destruction of a structure within a WUI perimeter. Structures within each wildfire event were grouped in order to prevent data leakage into the test set, and the 2 largest fires, Camp and Tubbs, were manually moved into the training set to allow the model to generalize as much as possible to unseen data for the model with all fires. In order to extract the best hyperparameters, grid-search on a chosen number of neurons per layer, dropout rate, learning rate, and regularization rate was done with cross-validation and early stopping based on PR-AUC for evaluation and predicted probability threshold selection. The best hyperparameters were then run on the test set with early stopping, resulting in the final Neural Network model used to create a metrics report, map visualizations, feature interpretations, and a structural loss report.

3 Results

3.1 Random Forest

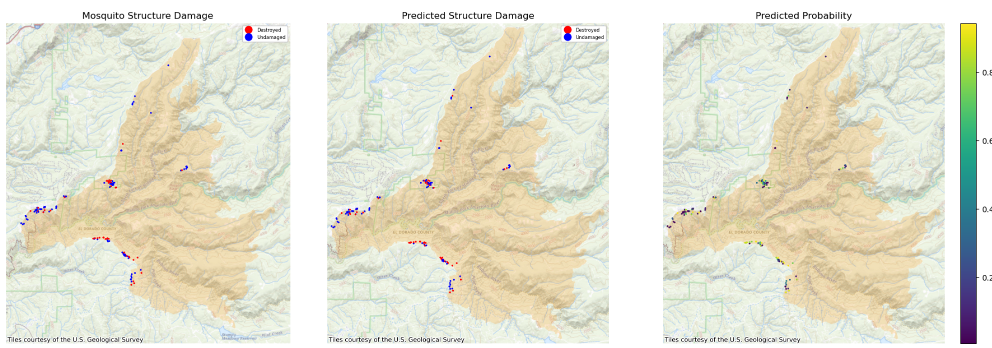

Figure 1: Observed and predicted structural damage for the model trained on all wildfire events. Points represent individual structures classified as undamaged or destroyed from the DINS dataset.

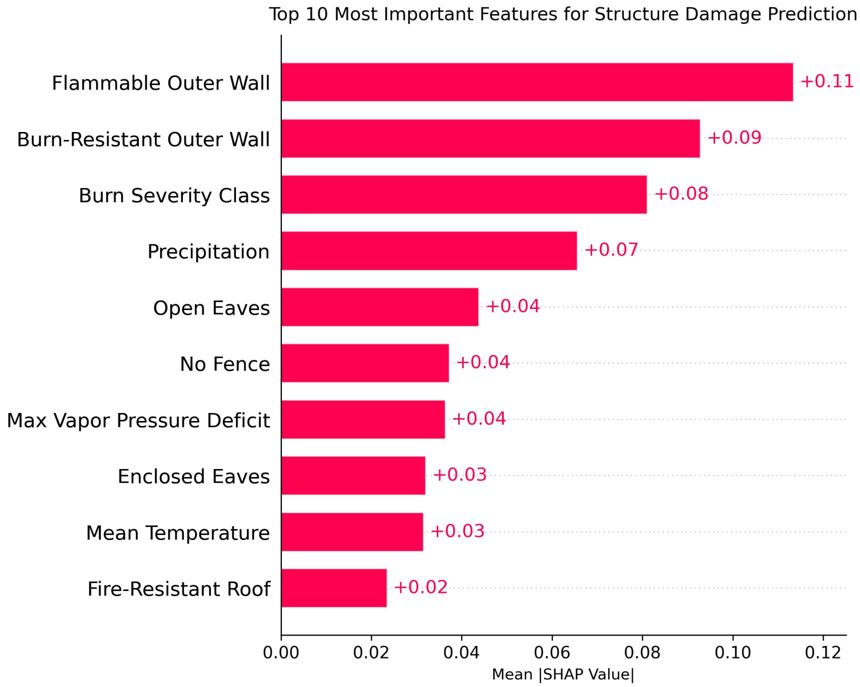

Figure 2: SHAP values detailing the top 10 most influential features for structure damage prediction for the Random Forest model trained on all fires.

Table 1: Classification report for structural damage prediction from Random Forest model trained on all fires.

Class	Precision	Recall	F1-score	Support
Undamaged	0.6420	0.8878	0.7452	303
Destroyed	0.9736	0.8930	0.9315	1402
Macro avg	0.8078	0.8904	0.8383	1705
Weighted avg	0.9146	0.8921	0.8984	1705
Accuracy	0.8921	–	–	1705

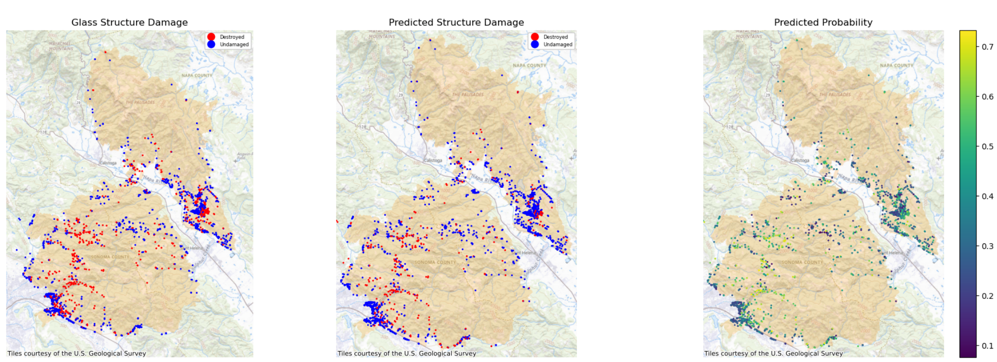

Figure 3: Observed and predicted structural damage for the Random Forest model trained on low-severity wildfire events. Points represent individual structures classified as undamaged or destroyed from the DINS dataset.

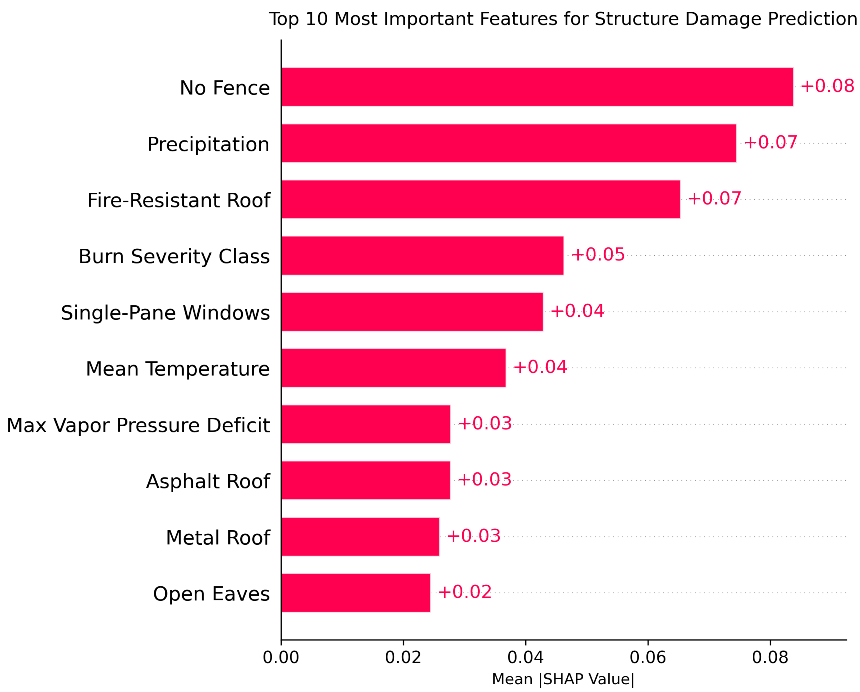

Figure 4: SHAP values detailing the top 10 most influential features for structure damage prediction for the Random Forest model trained on low-severity fires.

Table 2: Classification performance for wildfire structural damage prediction for the Random Forest model trained on low-severity fires.

Class	Precision	Recall	F1-score	Support
Undamaged	0.6453	0.9427	0.7661	2410
Destroyed	0.8636	0.4117	0.5576	2123
Macro avg	0.7545	0.6772	0.6619	4533
Weighted avg	0.7475	0.6940	0.6685	4533
Accuracy	0.6940	–	–	4533

The Random Forest model trained on all fires obtained a macro average F1-score of 0.83. The F1-score for destroyed structures is 0.93, while the F1-score for undamaged structures is 0.74. The top most influential features for this model include Flammable Outer Wall, Burn Resistant Outer wall, and Burn Severity Class.

The model trained on only low-severity fires obtained a macro average F1-score of 0.66, with an F1-score of 0.77 for undamaged structures and 0.56 for damaged structures. Among the most important features for this model are No Fence, Precipitation, and Fire-Resistant Roof.

Additional Random Forest models were trained for the remaining wildfire subsets, including high-severity fires, small fires, plume driven fires, and wind driven fires. Classification metrics, predicted damage maps, loss reports, confusion matrix heat maps, and SHAP feature importance graphs were generated for each model. The high-severity model had a macro average F1-score of 0.44, the small fires model had a macro average F1-score of 0.48, the plume driven model had a macro average F1-score of 0.58, and the wind driven model had macro average F1-score of 0.66.

3.2 Neural Network

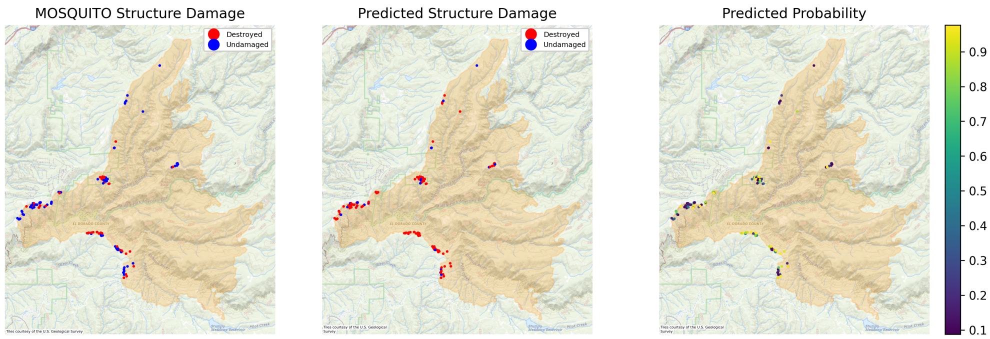

Figure 5: Observed and predicted structural damage for the model trained on all wildfire events. Points represent individual structures classified as undamaged or destroyed from the DINS dataset.

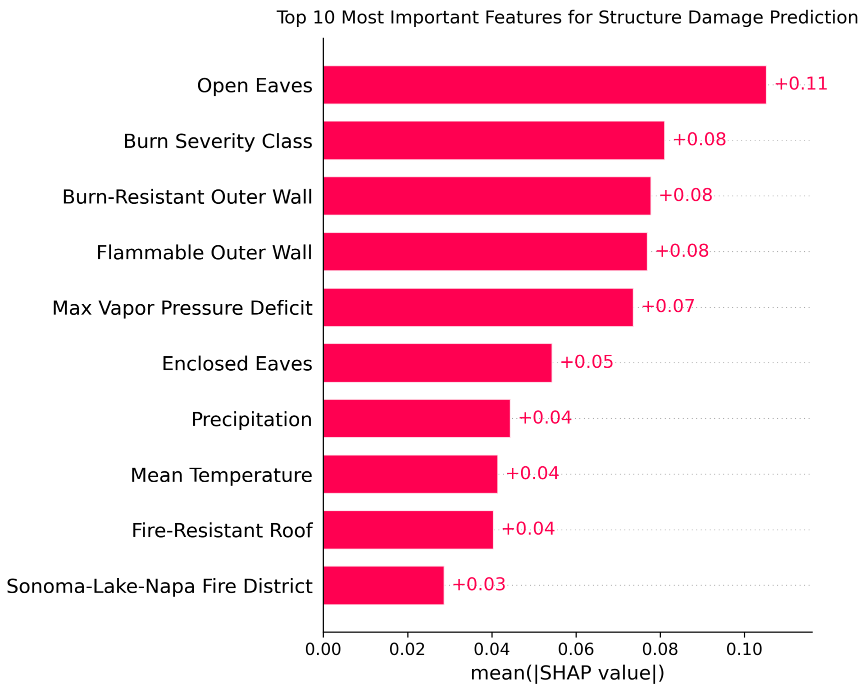

Figure 6: SHAP values detailing the top 10 most influential features for structure damage prediction for the Neural Network model trained on all fires.

Table 3: Classification report for structural damage prediction for the Neural Network model trained on all fires.

Class	Precision	Recall	F1-score	Support
Undamaged	0.8235	0.5083	0.6286	303
Destroyed	0.9018	0.9765	0.9377	1402
Macro avg	0.8627	0.7424	0.7831	1705
Weighted avg	0.8879	0.8933	0.8827	1705
Accuracy	0.8933	–	–	1705

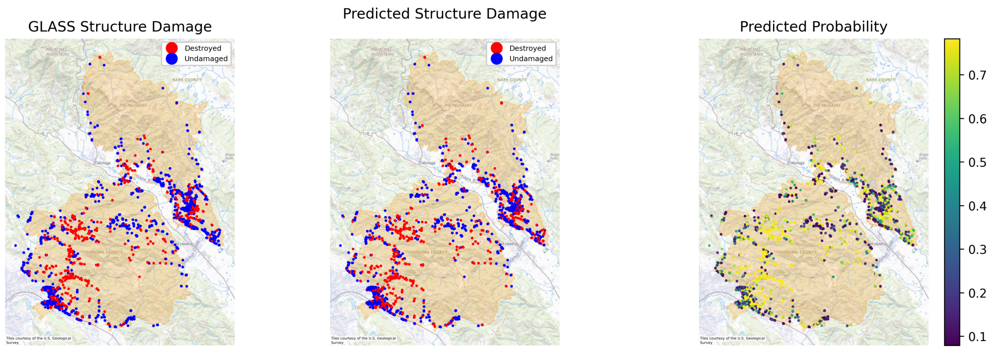

Figure 7: Observed and predicted structural damage for the Neural Network model trained on low-severity wildfire events. Points represent individual structures classified as undamaged or destroyed from the DINS dataset.

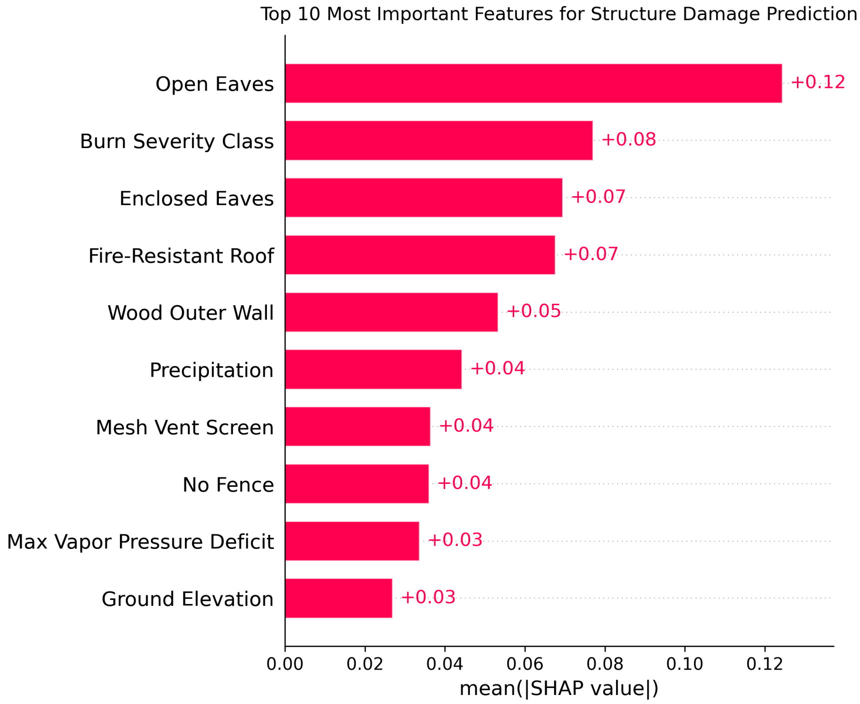

Figure 8: SHAP values detailing the top 10 most influential features for structure damage prediction for the Neural Network model trained on low-severity fires.

Table 4: Classification performance for wildfire structural damage prediction for Random Forest model trained on low-severity fires.

Class	Precision	Recall	F1-score	Support
Undamaged	0.8519	0.8925	0.8717	2410
Destroyed	0.8710	0.8238	0.8468	2123
Macro avg	0.8614	0.8582	0.8593	4533
Weighted avg	0.8608	0.8604	0.8600	4533
Accuracy	0.8604	–	–	4533

The Neural Network model trained on all fires obtained a macro average F1-score of 0.7831. The F1-score for destroyed structures is 0.9377, while the F1-score for undamaged structures is 0.6286. The top-most influential features for this model include Open Eaves, Burn Severity Class, and Flammable Outer Wall.

The model trained on only low-severity fires obtained a macro average F1-score of 0.8593, with an F1-score of 0.8717 for undamaged structures and 0.8468 for damaged structures. Among the most important features for this model are Open Eaves, Burn Severity Class, and Enclosed Eaves.

Additional Neural Network models were trained for the remaining wildfire subsets, including high-severity fires, small fires, plume driven fires, and wind driven fires. Classification metrics, predicted damage maps, loss reports, confusion matrix heat maps, and SHAP feature importance graphs were generated for each model. The high-severity model had a macro average F1-score of 0.4448, the small fires model had a macro average F1-score of 0.8085, the plume driven model had a macro average F1-score of 0.4584, and the wind driven model had macro average F1-score of 0.6740.

4 Discussion

The results from the Random Forest and Neural Network models demonstrate that residential structural damage in Wildfire-Urban Interface areas can be accurately and meaningfully predicted using publicly available geospatial, environmental, and structural data. Both models achieved strong performance when trained on all fires, with the Random Forest model achieving a macro F1-score of 0.84 and the Neural Network model achieving a macro F1-score of 0.78. This is a substantial improvement over the burn severity prediction models developed in Quarter 1, where our Neural Network performed best at roughly 40% accuracy. However, it's important to note that these results aren't directly comparable, as in Quarter 1, our task was a 4-class classification problem of predicting burn severity labels ranging from unburned to high burn. On the other hand, the Quarter 2 task is a binary classification problem that distinguishes destroyed structures from undamaged ones. Although it's likely that this difference in class size contributes directly to stronger performance metrics in Quarter 2, the significant improvement in the Quarter 2 task also likely displays a better alignment between our predictors and the target variable, as structure-level damage outcomes are more directly dependent on local environmental conditions and ground-level structural features than pixel-level burn severity classifications from satellite imagery.

Across both models, consistent sets of features emerged as the strongest and most influential predictors of structural damage. For both models trained on all fires, 9 out of the top 10 most influential features were shared between the Random Forest and Neural Network models, with the only difference being No Fence for the Random Forest and Sonoma-Lake-Napa Fire District for the Neural Network. This high similarity in top 10 predictors between two different models with distinctly different architectures strengthens our confidence that these features represent genuine predictors of structural damage and can't simply be attributed to model-specific variations. The shared top features include structural variables like Flammable and Burn-Resistant Outer Walls, Fire-Resistant Roofs, Burn Severity Class, and Open and Enclosed Eaves, as well as atmospheric variables like Precipitation, Mean Temperature, and Max Vapor Pressure Deficit. The prominence of outer wall material as a structural damage predictor is an intuitive finding, as structures with easily flammable outer walls are more likely to ignite when exposed to high heat or open flames. Similarly, it follows that structures with burn resistant outer walls would be found in areas of high wildfire risk, making it a strong predictor of structural damage likelihood. Additionally, open eaves are an expected predictor of wildfire risk, as they create easy pathways for wind-driven flames to enter a structure's interior and ignite. The presence of burn severity class reinforces that local fire intensity conditions play meaningful roles in determining damage outcomes on top of accounting for structural characteristics.

For both models trained on low severity fires, the two shared 6 of their top 10 features:

Burn Severity Class, Fire-Resistant Roof, Precipitation, Open Eaves, No Fence, and Max Vapor Pressure Deficit. However, open eaves showed a sharp divergence in relative importance for the two models, as it was the dominant predictor for the Neural Network at +0.12, whereas it was just a +0.02 for the Random Forest. This suggests that the Neural Network placed significantly more weight on open eaves enabling pathways for flames to enter structures and ignite. Furthermore, the models diverged on their remaining four features, as the Random Forest model identified Single-Pane Windows, Mean Temperature, Asphalt Roofs, and Metal Roofs as significant predictors, whereas the Neural Network highlighted Enclosed Eaves, Wood Outer Walls, Mesh Vent Screens, and Ground Elevation. These differences could likely be attributed to each model's sensitivity to different feature interactions and not a larger disagreement about the strongest drivers of damage risk, as both models still converge on many of the same top predictors. Despite these differences, we can observe a broad pattern across the low-severity models: outer wall material, which was among the strongest predictors in the all-fires models, drops considerably in prominence, while fire-resistant roof becomes noticeably more influential. Burn severity class remains a consistently strong predictor across all models. This shift suggests that in lower intensity fire conditions, where structures may be less likely to face overwhelmingly high heat levels, open eaves and roof construction become more decisive vulnerabilities. Atmospheric variables like precipitation and maximum vapor pressure deficit continue to have a considerable influence on damage prediction in both low severity models, reinforcing that weather conditions remain a meaningful contributor to damage outcomes regardless of fire intensity.

The models were trained on many fire subsets and identified important variations in predictive performance depending on fire type and severity. The Neural Network achieved a high macro F1-score of 0.86 when trained on the low severity fires, whereas the Random Forest model only achieved a macro F1-score of 0.67. This discrepancy might be highlighting the Neural Network's ability to better capture nonlinear interactions between features, especially in areas where fire intensity may not be the primary driver of structural damage outcomes. In contrast, both models performed poorly when trained on high severity fires, as the Neural Network had a macro average F1-score of 0.45 and the Random Forest model had a macro average F1-score of 0.44. This finding is not unsurprising, as in extreme fire conditions, structural characteristics may have less influence on damage outcomes, as nearly all structures in the fire's path would likely be destroyed regardless of their construction and materials, reducing the signal these models could learn from.

There are several important limitations to consider. First, the binary encoding of the target variable as either destroyed or undamaged excludes many structures with varying levels of damage in between our binary outcomes. This limits the ability of our models to produce nuanced risk assessment and could oversimplify the spectrum of potential wildfire damage outcomes. Second, there is notable class imbalance in our fire subsets, with damaged structures making up a significantly larger proportion of our training data than undamaged structures, as there were 1402 destroyed structures in our all-fires test set, compared to 303 undamaged structures. This class imbalance likely influenced our model behavior, as we can observe that the Random Forest model trained on all fires achieved a recall of 0.89 on undamaged structures, meaning it correctly identified most undamaged structures, but only achieved a precision of 0.64, meaning it also incorrectly classified many destroyed structures as undamaged. Despite using balanced class weights during hyperparameter tuning to address this, class imbalance couldn't be fully resolved. Third, even though data leakage was mitigated by grouping structures within fire events during cross-validation and manually assigned the two largest fires, Camp and Tubbs, to the training set, the generalizability of our models to different regions is uncertain, since our training data was limited to a specific set of California wildfires. It's yet to be determined if our models generalize well to new geographic regions that they haven't seen in their training or even future fires. Finally, the availability of data presented a limitation, as MTBS burn severity data, which provides essential fire perimeter data for our models, is typically released 1-2 years after a fire occurs due to the time required for satellite image processing. As a result, our analysis was limited to fires through 2023. This means that our models can't be quickly utilized on recent fires without waiting for updated MTBS releases. In the context of risk assessment, this constrains the applicability of our models for assessing damage from fires that have occurred recently.

Future work could expand this project in several directions. Expanding the target variable to multi-class damage instead of binary outcomes could enable nuanced outputs that are more actionable for residents and insurance companies. Incorporating more granular weather data such as wildfire direction, wind direction, and wind speed could improve predictions, especially in wind-driven fire scenarios, where our models underperformed. By collecting more subset-specific fire data to train our models, particularly in high severity and plume-driven wildfire scenarios, our models could be improved to generalize on areas where they underperformed. Additionally, replacing post-fire data dependencies with live weather feeds and updated structural inventories could move this framework toward real-time damage assessment during active fire events, aligning with our goal of developing accessible, cost-effective wildfire risk assessment using publicly available data. However, this would require significant refactoring of our data pipeline. Finally, this framework could be applied to fire-prone regions beyond California as a tool for wildfire risk assessment and community planning.

5 Conclusion

In this project, we investigated whether residential structure damage in Wildland-Urban Interface areas could be predicted using publicly available geospatial, environmental, and structural data. Motivated by the limited availability of structure-level wildfire risk assessment tools, we trained and evaluated Random Forest and Neural Network models on a comprehensive dataset integrating weather conditions, spatial features, burn severity information, and structural inspection data across 30 California wildfires.

Both models demonstrated that structure-level wildfire damage outcomes can be meaningfully predicted using publicly available data. When trained on all fires, the Random Forest model achieved a macro F1-score of 0.84 and the Neural Network achieved a macro F1-score of 0.78. Additionally, both models strongly agreed on their most influential features, as outer wall material, eave type, burn severity class, and atmospheric variables like precipitation and maximum vapor pressure deficit emerged as consistent predictors of structural damage across both models. Training models on specific fire subsets revealed that predictive performance varied greatly depending on fire type and severity, with low-severity fires producing strong results, while high-severity and plume-driven fires were more difficult to predict, likely due to the nature of extreme wildfire conditions burning indiscriminately, thereby reducing the predictive value of structural features.

These findings suggest that readily available data can serve as a practical and cost-effective foundation for structure-level wildfire damage assessment, which can offer insights to support better preparation for communities, as well as high-quality risk modeling for insurers in Wildland-Urban Interface areas. Future work extending this framework to multi-class damage outcomes and incorporating more data could strengthen the generalizability and practical impact of our project. As California continues to face increasing wildfire frequency and intensity due to climate change, accessible and easily interpretable tools for understanding and modeling structural vulnerability will become increasingly important.

References

CAL FIRE. 2022. "CAL FIRE Damage Inspection (DINS) Data." https://hub-calfire-forestry.hub.arcgis.com/datasets/cal-fire-damage-inspection-dins-data/about

Inventory, National Structure. n.d.. "Structure Data." https://nsi.sec.usace.army.mil/downloads/

Kramer, H. Anu, Miranda H. Mockrin, Patricia M. Alexandre, Susan I. Stewart, and Volker C. Radeloff. 2018. "Where wildfires destroy buildings in the US relative to the wildland–urban interface and national fire outreach programs." *International Journal of Wildland Fire* 27 (5), p. 329–341. [Link]

LANDFIRE. 2025. "LANDFIRE WCS/WMS." https://landfire.gov/data/lf_wcs_wms

Mahmoud, Hussam, and Akshat Chulahwat. 2020. "Assessing wildland-urban interface fire risk." https://doi.org/10.1098/rsos.201183

MTBS. 2025. "Direct Download MTBS." https://www.mtbs.gov/direct-download

Multi-Resolution Land Characteristics Consortium. 2025. "MRLC Data." https://www.mrlc.gov/data

PRISM Climate Group, Oregon State University. 2025. "PRISM Weather Data." https://prism.oregonstate.edu/

SILVIS Lab. n.d.. "Wildland-Urban Interface." https://silvis.forest.wisc.edu/maps-data/wildland-urban-interface/

Appendices

A.1 Training Details

A.1.1 Random Forest

The final random forest model was selected based on the best hyperparameters and threshold. The hyperparameters were selected using a 3 fold cross validation. The parameters tuned were:

$$n\_estimators \in \{300, 500, 800\}$$

$$max\_depth \in \{20, 40, 60, None\}$$

$$min\_samples\_split \in \{2, 5, 10\}$$

$$min\_samples\_leaf \in \{1, 2, 5\}$$

$$max\_features \in \{"sqrt", 0.3, 0.5\}$$

Then, the threshold was tuned from a range of 0.2 to 0.8 using the validation set to determine the best probability threshold cutoff.

A.1.2 Neural Network

The final trained neural network was selected based on the best architecture and hyperparameters chosen from the 5-fold cross validation on the combined training and validation sets. The number of units for each fully connected layer was n_units = [128, 64, 32, 16, 8], activation was set to "relu", dropout_rate to 0.15, use_batchnorm to "True", learning_rate to 0.01, and to l2_reg to 0.001. Early stopping of 15 epochs was used on the validation set with the best epoch step with the best validation accuracy being extracted as the final model. In order to get the feature importances, SHAP values were computed using the "shap" package and used to create the feature importance visualizations for the report.