Introduction

An earthquake is what happens when two blocks of the earth suddenly slip past one another. The surface where they slip is called the fault or fault plane. The location below the earth’s surface where the earthquake starts is called the hypocenter, and the location directly above it on the surface of the earth is called the epicenter. Gorkha earthquake occurred at 11:56 Nepal Standard Time on 25 April 2015, with a magnitude of 8.1Ms. Its epicenter was east of Gorkha District at Barpak, Gorkha, and its hypocenter was at a depth of approximately 8.2. It represents the worst natural disaster to strike Nepal since the 1934 Nepal-Bihar earthquake. The earthquake triggered an avalanche on Mount Everest, killing 21, making 25 April 2015 the deadliest day on the mountain in history. The earthquake triggered another huge avalanche in the Langtang valley, where 250 people were reported missing. Hundreds of thousands of Nepalese were made homeless with entire villages flattened, across many districts of the country. Centuries-old buildings were destroyed at UNESCO World Heritage Sites in the Kathmandu Valley, including some at the Kathmandu Durbar Square, the Patan Durbar Square, the Bhaktapur Durbar Square, the Changu Narayan Temple, the Boudhanath stupa and the Swayambhunath Stupa. Each year, thousands of earthquakes occur throughout the world. Although not damaging, few earthquakes provide a wealth of information that enables seismologists and engineers to better assess the distribution, frequency, and severity of seismic hazards throughout the country. Seismograph networks supply earthquake parameter and waveform data that are essential for the real-time evaluation of tectonic activity for public safety, the development of earthquake hazard maps and seismic design criteria used in building codes and land-use planning decisions.

**Nepal**, [country](https://www.britannica.com/topic/nation-state) of [Asia](https://www.britannica.com/place/Asia), lying along the southern slopes of the [Himalayan mountain ranges](https://www.britannica.com/place/Himalayas). It is a landlocked country located between [India](https://www.britannica.com/place/India) to the east, south, and west and the [Tibet Autonomous Region of China](https://www.britannica.com/place/Tibet) to the north. Its territory extends roughly 500 miles (800 kilometres) from east to west and 90 to 150 miles from north to south. The capital is [Kathmandu](https://www.britannica.com/place/Kathmandu).

Nepal contains some of the most rugged and difficult mountain terrain in the world. Roughly 75 percent of the country is covered by mountains. From the south to the north, Nepal can be divided into four main physical belts, each of which extends east to west across the country. These are, first, the [Tarai](https://www.britannica.com/place/Tarai), a low, flat, fertile land [adjacent](https://www.merriam-webster.com/dictionary/adjacent) to the border of India; second, the forested Churia foothills and the Inner Tarai zone, rising from the Tarai plain to the rugged Mahābhārat Range; third, the mid-mountain region between the Mahābhārat Range and the Great Himalayas; and, fourth, the [Great Himalaya Range](https://www.britannica.com/place/Great-Himalayas), rising to more than 29,000 feet (some 8,850 metres).

Wedged between two giants, India and China, Nepal seeks to keep a balance between the two countries in its foreign policy—and thus to remain independent. A factor that contributes immensely to the geopolitical importance of the country is the fact that a strong Nepal can [deny](https://www.britannica.com/dictionary/deny) China access to the rich Gangetic Plain; Nepal thus marks the southern boundary of the Chinese sphere north of the Himalayas in Asia.

Objective:

Our Motive is to devise a model which helps the Building constructors and related workers to construct the building thereby the building can withstand mega quakes in future. Based on information collected from previous earthquakes and the damages it caused we can coin damage levels a building can face when a similar earthquake strikes again.

As the name suggests, the project involves predict and prevent earthquake damages to an extent unlike specifically damage from the [Gorkha earthquake](https://en.wikipedia.org/wiki/April_2015_Nepal_earthquake) which occurred in April 2015 and killed over 9,000 people.

Our task in this project to forecast how badly an individual house is damaged, given the information about its location, secondary usage, and the materials used to build the house in the first place. The damage grade of each house is stated as an integer variable between one and three.

Few advantages of these kind of predictions are implementing precautionary measures and

better disaster preparedness like:

▪ Devise a better a building plan with respect to area

▪ Prioritizing Renovation of old/damage prone buildings.

▪ Classification of geographic in which the building exists.

▪ Helps in construction of Building structure in any area.

▪ Approving Number of floors while construction.

Data:

The main data taken from Data Driven

The dataset mainly consists of information on the buildings' structure and their legal ownership. Each row in the dataset represents a specific building in the region that was hit by Gorkha

earthquake. There are 200k observations and 39 columns in this dataset, where the building\_id

column is a unique and random identifier. The remaining 38 features are described in the Data

preprocessing section.

Based on aspects of building location and construction, our goal is to predict the level of

damage to buildings caused by the 2015 Gorkha earthquake. The data was collected through

household surveys using mobile technology by Kathmandu Living Labs and the Central

Bureau of Statistics, which works under the National Planning Commission Secretariat of

Nepal in the earthquake-affected districts.

This survey is one of the largest post-disaster datasets ever collected, containing valuable

information on earthquake impacts, household conditions, and socio-economic-demographic

statistics.

Below is the description of the variables:

• geo\_level\_1\_id, geo\_level\_2\_id, geo\_level\_3\_id (type: int): geographic region in

which building exists, from largest (level 1) to most specific sub-region (level 3).

Possible values: level 1: 0-30, level 2: 0-1427, level 3: 0-12567.

• count\_floors\_pre\_eq (type: int): number of floors in the building before the

earthquake.

• age (type: int): age of the building in years.

• area\_percentage (type: int): normalized area of the building footprint.

• height\_percentage (type: int): normalized height of the building footprint.

• land\_surface\_condition (type: categorical): surface condition of the land where the

building was built. Possible values: n, o, t.

• foundation\_type (type: categorical): type of foundation used while building. Possible

values: h, i, r, u, w.

• roof\_type (type: categorical): type of roof used while building. Possible values: n, q,

x.

STAT 642-675 GROUP 6 SPRING 2020

3

• ground\_floor\_type (type: categorical): type of the ground floor. Possible values: f,

m, v, x, z.

• other\_floor\_type (type: categorical): type of constructions used in higher than the

ground floors (except of roof). Possible values: j, q, s, x.

• position (type: categorical): position of the building. Possible values: j, o, s, t.

• plan\_configuration (type: categorical): building plan configuration. Possible values:

a, c, d, f, m, n, o, q, s, u.

• has\_superstructure\_adobe\_mud (type: binary): flag variable that indicates if the

superstructure was made of Adobe/Mud.

• has\_superstructure\_mud\_mortar\_stone (type: binary): flag variable that indicates

if the superstructure was made of Mud Mortar - Stone.

• has\_superstructure\_stone\_flag (type: binary): flag variable that indicates if the

superstructure was made of Stone.

• has\_superstructure\_cement\_mortar\_stone (type: binary): flag variable that

indicates if the superstructure was made of Cement Mortar - Stone.

• has\_superstructure\_mud\_mortar\_brick (type: binary): flag variable that indicates

if the superstructure was made of Mud Mortar - Brick.

• has\_superstructure\_cement\_mortar\_brick (type: binary): flag variable that

indicates if the superstructure was made of Cement Mortar - Brick.

• has\_superstructure\_timber (type: binary): flag variable that indicates if the

superstructure was made of Timber.

• has\_superstructure\_bamboo (type: binary): flag variable that indicates if the

superstructure was made of Bamboo.

• has\_superstructure\_rc\_non\_engineered (type: binary): flag variable that indicates

if the superstructure was made of non-engineered reinforced concrete.

• has\_superstructure\_rc\_engineered (type: binary): flag variable that indicates if the

superstructure was made of engineered reinforced concrete.

• has\_superstructure\_other (type: binary): flag variable that indicates if the

superstructure was made of any other material.

• legal\_ownership\_status (type: categorical): legal ownership status of the land where

building was built. Possible values: a, r, v, w.

• count\_families (type: int): number of families that live in the building.

• has\_secondary\_use (type: binary): flag variable that indicates if the building was

used for any secondary purpose.

• has\_secondary\_use\_agriculture (type: binary): flag variable that indicates if the

building was used for agricultural purposes.

• has\_secondary\_use\_hotel (type: binary): flag variable that indicates if the building

was used as a hotel.

• has\_secondary\_use\_rental (type: binary): flag variable that indicates if the building

was used for rental purposes.

• has\_secondary\_use\_institution (type: binary): flag variable that indicates if the

building was used as a location of any institution.

• has\_secondary\_use\_school (type: binary): flag variable that indicates if the building

was used as a school.

• has\_secondary\_use\_industry (type: binary): flag variable that indicates if the

building was used for industrial purposes.

• has\_secondary\_use\_health\_post (type: binary): flag variable that indicates if the

building was used as a health post.

• has\_secondary\_use\_gov\_office (type: binary): flag variable that indicates if the

building was used as a government office.

• has\_secondary\_use\_use\_police (type: binary): flag variable that indicates if the

building was used as a police station.

• has\_secondary\_use\_other (type: binary): flag variable that indicates if the building

was secondarily used for other purposes.

We are going to predict damage\_grade class, which represents a level of damage to

the building that was hit by the earthquake. There are 3 grades/classes of the damage:

▪ 1 represents low damage

▪ 2 represents a medium amount of damage

▪ 3 represents almost complete destructio