

Flood Risk Index ML

Common causes of floods:

- Heavy rainfalls \Rightarrow overflowing rivers.
- Quick melting of snowpacks
- Human-influenced causes:
 - \hookrightarrow Structural failure of dams (catastrophic flooding)
 - \hookrightarrow Building of impermeable surfaces (roads) \rightarrow prevents soaking into ground.
 - \hookrightarrow Climate change \rightarrow rise of ocean levels.
- Ice jams (mostly in Canada) \rightarrow form dams on rivers.
- Coastal storms \rightarrow push large amounts of ocean water on coast.

Parts of the world that are affected:

- Southeast Asia
 - \hookrightarrow Bangladesh \Rightarrow top 2 country w/ % population at risk (57.5%)
 - \hookrightarrow India and China = largest populations at risk.
 - \hookrightarrow concentration of population in low-lying coastal areas.
- America
 - \hookrightarrow highest nb. of significant floods in 2023.
 - \hookrightarrow coastal regions like California / Florida
- Africa
 - \hookrightarrow Egypt is in the top five for population exposure to flood risk.
- Europe
 - \hookrightarrow Netherlands \Rightarrow top 1 country w/ % population at risk (58.7%)

Indonesia

- According to BNPB, floods are the most common type of natural disaster in Indonesia
- Causes:
 - \hookrightarrow Monsoon season
 - \hookrightarrow Low altitude and coastal cities (altitude of Jakarta is 4m)
 - \hookrightarrow Indonesian rivers overflow easily.
 - \hookrightarrow Poor drainage in some regions.

Data Needed

- * Meteorological data → precipitation ✓
 - ↳ temperature (for evaporation / snowmelt) ✓
 - ↳ humidity (for storm severity) ✓
- * Hydrological data → real-time data on river levels / flow rates
 - ↳ soil moisture (saturation level of the soil)
 - ↳ amount of snow on the ground.
- * Land features → low elevation floods first
 - ↳ slope (flat areas accumulate water easily)
 - ↳ closeness of city to river/water areas.
 - ↳ land ??? (concrete vs. vegetation).

* Targets → Flood Risk Index : range 0 to 1.

The target must be computed using all the data and factors stated above. Each data will have a different weight, so must use standard scaling to make sure less relevant data does not influence prediction too much.

- * Raw data:
 - one CSV file per city
 - each row = one day from 01/01/2015 to 01/12/2025
 - each column = one feature (rainfall, temp, etc.)
 - targets = real data or computed?