

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penurunan muka tanah merupakan fenomena lingkungan yang menyebabkan permukaan bumi mengalami penurunan secara bertahap atau tiba-tiba (Widodo et al., 2022). Penurunan muka tanah telah terjadi di DKI Jakarta yang disebabkan oleh berbagai faktor seperti eksploitasi air tanah yang berlebihan, beban infrastruktur, dan kondisi geologi (Bui et al., 2018). Untuk pemantauan penurunan tanah, terdapat teknologi Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) yang banyak digunakan, PS-InSAR mendeteksi perubahan permukaan tanah melalui analisis data radar multi-temporal (Bui et al., 2018). Akurat dalam analisis historis, PS-InSAR memiliki keterbatasan dalam memprediksi dan tidak dapat memodelkan pola penurunan muka tanah untuk masa depan. Untuk mengatasi keterbatasan tersebut, pendekatan dengan machine learning seperti Random Forest penting untuk memprediksi pola penurunan tanah berdasarkan data multivariat.

Penurunan muka tanah diprediksi menggunakan algoritma *Bayesian Logistic Regression* (BLR), *Support Vector Machine* (SVM), *Random Forest* (RF). Dengan ketiga algoritma tersebut, *Random Forest* sebagai salah satu jenis machine learning yang mampu mengurangi noise dan mengurangi dampak overfitting dengan teknik ensemble. Akurasi data cukup unggul pada data linear maupun non-linear dibandingkan dengan SVM kurang optimal dan LBR hanya efisien untuk dataset kecil. Penurunan muka tanah sendiri memiliki data yang besar dan kompleks, random forest mampu menangani dibandingkan SVM dan BLR. Meskipun ketiga metode tersebut memiliki keunggulan satu sama lain, berdasarkan dengan data penelitian, data berupa data besar dan kompleks memerlukan random forest untuk mengurangi adanya overfitting dibandingkan yang lain (Bui et al., 2018; Ahmed et al., 2024; Hosseinzadeh et al., 2024; Liu & Rahul, 2024; Handika et al., 2024).

Pada penelitian sebelumnya, Ahmed et al., (2024) telah membandingkan algoritma Random Forest dengan LightGBM dan XGBoost. Hasil yang diperoleh menunjukkan random forest lebih baik dibandingkan dengan LightGBM dan XGBoost dengan nilai ($R^2 = 0,84$), ($KGE = 0,89$), ($RMSE = 2,19$), dan ($MAE = 1,42$). Pada penelitian Handika et al., (2024) kinerja model Random Forest memiliki hasil yang lebih baik dengan nilai 0,73 dibandingkan dengan SVM. Terdapat kekurangan pada random forest apabila tuning parameter buruk akan meningkatkan bias maka akan peneliti perbaiki dengan penyetelan hyperparameter. Untuk mengatasi keterbatasan ini, pada penelitian Ahmed et al., (2024) random forest sebagai algoritma prediksi penurunan tanah menggunakan metrix evaluasi *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), *Coefficient Determination* (R^2) dan *Kling-Gupta Efficiency* (KGE) untuk memprediksi pola penurunan tanah. Dan dalam mengatasi tuning parameter buruk dilakukan penyetelan hyperparameter. Penyetelan hyperparameter untuk mengoptimalkan parameter dengan menggunakan *Randomized SearchCV* yang dapat mengurangi rentang pencarian hyperparameter dan *Grid SearchCV* untuk menentukan set hyperparameter terbaik.

Peneliti akan menggunakan empat matrix untuk menilai akurasi model menggunakan *Mean Absolute Error* (MAE) mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai actual. MAE menilai prediksi dari nilai sebenarnya dalam satuan yang sama dengan data. *Root Mean Square Error* (RMSE) menghitung akar kuadrat dari rata-rata kesalahan kuadrat. RMSE memberikan Gambaran tentang kesalahan prediksi dengan memberi bobot lebih besar pada kesalahan yang lebih besar. *Coefficient of Determination* (R^2) mengukur seberapa baik model dapat menjelaskan variabilitas data. Nilai R^2 berada antara 0 hingga 1 dengan nilai mendekati 1 dengan menunjukan model dikatakan baik. *Kling-Gupta Efficiency* (KGE) dirancang untuk mengatasi kelemahan dalam evaluasi tradisional (R^2). KGE mempertimbangkan korelasi, bias relative serta rasio varians.

Penggunaan hyperparameter ini menyesuaikan pohon keputusan (*n_estimators*, *max_features*, dan *min_samples_leaf*) dengan menggunakan timeout pekerja (*joblib*) untuk memastikan pekerja memiliki cukup waktu serta threading agar tidak membutuhkan CPU yang banyak. Dengan demikian, Random Forest diharapkan dapat memprediksi penurunan muka tanah dengan kondisi data yang besar dan kompleks dan diharapkan tuning parameter dapat teratasi dengan baik. Sehingga penelitian ini dapat dibandingkan antara metode machine learning dengan SAR.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, terdapat beberapa masalah utama yang diidentifikasi dalam penelitian ini, yaitu:

1. Penurunan tanah merupakan bahaya utama bagi stabilitas permukaan tanah di DKI Jakarta. Penurunan tanah terjadi di beberapa wilayah DKI Jakarta khususnya di Pantai Indah Kapuk, Marunda, Ancol, Jakarta Utara, dan Kembangan, Jakarta Barat. Di sisi lain, penurunan tanah yang sedikit terjadi di wilayah Kalibaru, Jakarta Pusat.
2. Penurunan lebih signifikan terjadi di wilayah reklamasi seperti Pantai Indah Kapuk dibandingkan dengan karakteristik tanah aluvial.
3. Metode survei lapangan konvensional dengan teknologi Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) sudah dilakukan namun ingin membandingkan hasilnya dengan pendekatan machine learning khususnya random Forest terutama di wilayah DKI Jakarta berdasarkan data yang sudah.

1.3 Tujuan Masalah

1. Menganalisis penurunan tanah yang ada di DKI Jakarta serta penurunan tanah periode 2017 – 2021 dari tahun ke tahun.
2. Menganalisis wilayah reklamasi dengan wilayah yang memiliki karakteristik tanah aluvial di DKI Jakarta.
3. Hasil analisis menggunakan random forest akan dibandingkan dengan hasil sebelumnya menggunakan PS-InSAR.

4. Memprediksi penurunan tahun 2022 menggunakan dataset tahun 2017-2021

1.4 Batasan Masalah

Batasan masalah penelitian ini dibuat untuk membuat penelitian lebih fokus dan mencapai tujuan:

1. Penelitian ini menggunakan data yang sudah ada sebelumnya tentang penurunan muka tanah dan data yang digunakan dibatasi pada periode waktu tertentu yaitu tahun 2017 - 2021.
2. Dalam penelitian ini, penggunaan algoritma machine learning dibatasi pada Random Forest untuk memprediksi penurunan muka tanah yang ada di DKI Jakarta sehingga penelitian ini hanya mengeksplorasi satu algoritma tersebut.
3. Fokus utama penelitian ini membandingkan hasil analisis dari dua metode yang berbeda yaitu PS-InSAR dan Random Forest.
4. Penelitian ini memiliki batasan wilayah yang memprediksi pada satu provinsi saja yaitu provinsi DKI Jakarta.

1.5 Kontribusi

Pada penelitian ini diharapkan dapat memberikan sumbangan baik:

1. Secara teoritis, penelitian ini memberikan peluang terkait teori, apakah machine learning dengan algoritma Random Forest dapat menganalisis penurunan muka tanah dengan hasil yang optimal.
2. Secara praktis, hasil penelitian ini memberikan hasil yang dapat membantu membandingkan hasil analisis antara dua metode yang berbeda.