

## 1. Pendahuluan

### Latar Belakang

Saham merupakan instrumen keuangan pasar modal yang menarik perhatian banyak orang. Dengan melakukan investasi pada suatu saham seorang pelaku investasi dapat mendapatkan keuntungan yang cukup besar dengan usaha yang sedikit namun tidak selamanya melakukan investasi dapat mendatangkan keuntungan. Dengan berkembangnya *social media*, banyak masyarakat terjun ke dalam dunia investasi tanpa memiliki ilmu yang memadai [1] dan pada akhirnya masyarakat tersebut malah mengalami kerugian. Hal tersebut dapat diartikan bahwa diperlukannya ilmu yang cukup dalam melakukan analisis *return* saham. Maka dari itu telah banyak dilakukan penelitian dalam mempelajari fluktuasi naik turunnya harga saham sehingga keuntungan yang didapatkan semakin meningkat dan tentunya menghindari kerugian.

Metode analisis yang digunakan dalam mempelajari fluktuasi naik turunnya harga saham tersebut terbagi menjadi 2 yaitu *time-series* analisis dan *cross-sectional*. *Time-Series* menggunakan data deret waktu yang merupakan suatu representasi dari variabel dalam interval waktu tertentu. Data tersebut akan dianalisis dan dipelajari polanya sehingga dapat memprediksi nilai di masa depan. Analisis *cross-sectional* digunakan untuk membandingkan bermacam-macam variabel yang diambil pada suatu periode waktu tertentu untuk menentukan pengaruh dari masing-masing variabel terhadap suatu hasil yang dicari [2].

Analisis *time-series* merupakan metode paling banyak digunakan karena telah memberikan hasil prediksi yang cukup baik dalam beberapa tahun kebelakang, berbeda halnya dengan analisis *cross-sectional* yang masih jarang digunakan namun diyakini dapat memberikan hasil prediksi yang lebih baik. Pada metode *cross-sectional* terdapat faktor yang dikatakan memiliki tingkat prediksi yang tinggi. Maka dari itu *cross-sectional* tidak hanya menggunakan data historis saja, metode ini juga menggunakan data fundamental sebagai faktor-faktor pertimbangan dalam memprediksi *return* saham.

Pada tugas akhir ini membangun model prediksi *return* saham menggunakan *cross-sectional machine learning* pada saham yang tergabung dalam indeks LQ45 dan membandingkan performansi *cross-sectional machine learning* dengan *time-series machine learning*. Saham dengan *return* saham yang melebihi *threshold* yang ditentukan akan dipilih dalam pembentukan portofolio.

### Topik dan Batasannya

Topik yang akan dibahas pada tugas akhir ini adalah implementasi *cross-sectional machine learning* dengan algoritma *Random Forest* dalam memprediksi *return* saham yang tergabung dalam indeks LQ45 serta membangun portofolio berdasarkan prediksi *return* saham tersebut. Batasan masalah yang digunakan untuk data teknis merupakan data historis penutupan harga harian saham indeks LQ45 yang didapatkan dari Yahoo Finance sedangkan untuk data fundamental dihitung setiap kuartal dan di dapatkan dari IDX. Rentang waktu data yang digunakan pada tugas akhir ini yaitu 28 April 2016 hingga 30 Juni 2022.

### Tujuan

Tujuan tugas akhir ini adalah mengimplementasikan metode *cross-sectional machine learning* dengan algoritma *Random Forest* dalam memprediksi *return* saham yang tergabung pada indeks LQ45 dan membandingkannya dengan *time-series machine learning*. Pengaplikasian dari kedua buah model prediksi tersebut akan digunakan dalam seleksi portofolio.

### Organisasi Tulisan

Organisasi tulisan pada tugas akhir ini setelah pendahuluan yaitu studi terkait mengenai analisis *cross-sectional*, *Random Forest*, dan konstruksi portofolio.

