
Analisis dan Prediksi Kinerja Saham e-IPO Menggunakan Pendekatan Tabular Transformer dan *Attention Mechanism*

Elena Angkawi

Department of Computer Science
Bina Nusantara University
Jakarta, DKI Jakarta, 11480
elena.angkawi@binus.ac.id

Filip Nathan

Department of Computer Science
Bina Nusantara University
Jakarta, DKI Jakarta, 11480
filip.nathan@binus.ac.id

William Wijaya

Department of Computer Science
Jakarta, DKI Jakarta, 11480
william.wijaya028@binus.ac.id

Abstrak

Electronic Indonesia Public Offering (IPO) adalah metode perusahaan untuk memperoleh pendanaan dengan menjual saham kepada publik. Sejak tahun 2021, di Indonesia telah diterapkan sistem e-IPO sebagai mekanisme digital yang bertujuan untuk mempermudah proses penawaran saham perdana secara elektronik dan meningkatkan transparansi serta aksesibilitas bagi investor ritel. Penelitian ini menggunakan pendekatan Tabular Transformer berbasis attention mechanism, yaitu TabPFN dan TabNet, serta akan dibandingkan dengan model yang konvensional seperti XGBoost. Temuan menunjukkan bahwa TabPFN dan TabNet memberikan akurasi lebih tinggi dan prediksi yang lebih stabil, dengan evaluasi menggunakan *F1-Score (weighted average)* mencapai 0,72. Adanya *attention mechanism* memungkinkan model untuk memfokuskan perhatian kepada fitur-fitur penting instrument saham, sehingga lebih efektif dalam menganalisis performa e-IPO dari setiap emiten. Proses *Exploratory Data Analysis (EDA)* menunjukkan bahwa variabel seperti kuartal listing, indeks VIX, dan metrik keuangan perusahaan tidak memiliki pengaruh signifikan terhadap performa saham e-IPO. Sebaliknya, faktor seperti tipe underwriter, total aset perusahaan, dan papan pencatatan justru menunjukkan korelasi yang lebih kuat dengan kinerja saham setelah e-IPO. Berdasarkan analisis LIME (*Local Interpretable Model-agnostic Explanations*), keberhasilan prediksi performa e-IPO menggunakan fundamental perusahaan dan reputasi underwriter. Pendekatan ini dapat menjadi alat bantu pengambilan keputusan investasi yang cerdas dan berdasarkan data di pasar modal Indonesia.

Kata kunci: e-IPO, TabPFN, TabNet, *attention mechanism*, LIME

1 Pendahuluan

1.1 Latar Belakang

Initial Public Offering (IPO) adalah salah satu cara bagi perusahaan untuk mendapatkan pendanaan melalui pasar modal dengan melepas sebagian kepemilikannya kepada publik. Melalui IPO, perusahaan akan menjadi PT Terbuka (*go public*) yang di mana data keuangannya dapat diakses oleh publik pada tempat sahamnya terdaftar. Dana yang diperoleh umumnya dimanfaatkan perusahaan untuk berbagai keperluan seperti ekspansi bisnis, akuisisi, biaya modal, pembayaran hutang, dan lain-lain (Amor and Kooli, 2017). Per akhir tahun 2024, terdapat 943 total emiten pada Bursa Efek Indonesia, dengan mayoritas tercatat melalui IPO [1].

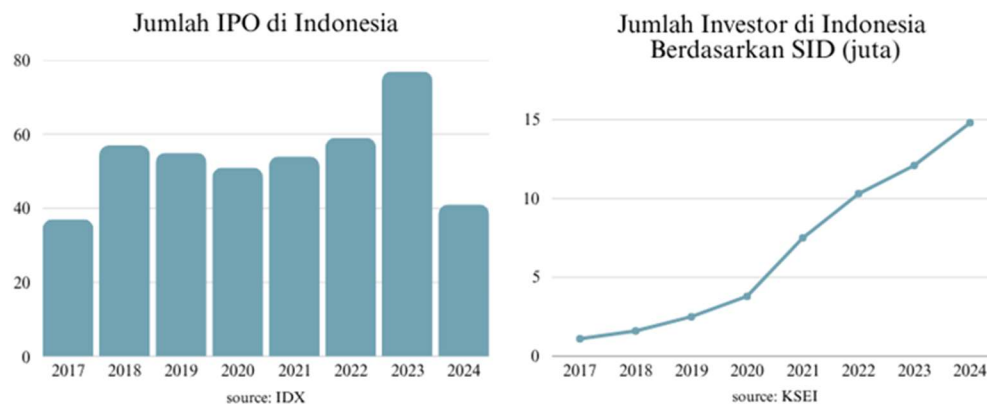


Figure 1: (a) Data Jumlah IPO [1]. (b) Data Jumlah Investor [2].

Dengan jumlah investor pasar modal yang terus bertumbuh secara signifikan, dibutuhkan sistem informasi yang mendukung investor untuk berinvestasi dengan lebih mudah dan transparan [3]. Untuk menjawab itu, Otoritas Jasa Keuangan (OJK) pada Desember 2020 menerbitkan regulasi OJK nomor 41/POJK.04/2020 mengenai implementasi *Electronic Public Offerings of Equity, Debt Securities, and/or Sukuk* ("POJK 41") [4]. Sejak diberlakukannya POJK 41, seluruh proses *book building*, *ordering*, dan *allotment* dari sekuritas dilakukan menggunakan sistem informasi digital yang dinamakan *Electronic Indonesia Public Offering* (e-IPO). Sistem e-IPO diharapkan dapat menciptakan proses *book building* yang lebih efisien dan transparan, serta mendorong partisipasi investor yang lebih luas dalam pasar modal [5].

Perkembangan dari IPO yang meningkat setiap tahunnya didampingi dengan bertambahnya jumlah investor yang mengikuti e-IPO [3]. Hal ini menunjukkan keberhasilan sistem e-IPO dalam meningkatkan partisipasi para pelaku pasar modal. Namun demikian, peningkatan ini dibarengi munculnya perilaku-perilaku investor yang cenderung spekulatif dan kurang rasional. Salah satu contohnya adalah *herding*, yaitu keputusan investasi yang dipengaruhi oleh tindakan kolektif, bukan berdasarkan analisis fundamental [6]. Selain itu, terdapat pula pola pikir untuk memperoleh keuntungan cepat dari potensi kenaikan harga saham pada hari pertama perdagangan, biasanya disertai dengan *fear of missing out* (FOMO) [7]. Kedua perilaku ini sering kali mendorong tindakan spekulatif yang tidak diiringi oleh pemahaman mendalam terhadap prospektus emiten atau kondisi keuangannya.

Tidak semua saham yang melakukan IPO berakhir dengan performa yang baik di pasar sekunder. Meski demikian, kinerja saham yang buruk tidak selalu mencerminkan kondisi fundamental perusahaan yang lemah. Berdasarkan data dari IDX, tercatat sekitar 167 saham dari 906 saham tercatat memiliki harga penutupan sebesar Rp50 atau lebih rendah, yang mencakup sekitar 18% dari total saham tercatat. Data ini telah mengecualikan saham yang berada dalam pengawasan khusus, tidak likuid, atau tengah disuspensi (IDX, 2024). Mekanisme pasar yang kompleks, termasuk faktor psikologis investor, ekspektasi jangka pendek, serta informasi yang asimetris, turut memengaruhi

performa saham pasca-IPO. Fenomena ini menunjukkan bahwa pergerakan harga saham tidak selalu didorong oleh nilai intrinsik perusahaan, melainkan lebih sering dipengaruhi oleh tren pasar dan sentimen jangka pendek (Michael Yeoh, 2025).

Untuk memahami faktor-faktor yang memengaruhi kinerja saham e-IPO, diperlukan pendekatan analitik yang mampu mengungkap pola tersembunyi, khususnya di tengah keterbatasan data historis yang tersedia. Perkembangan model *artificial intelligence* berbasis *transformer* memungkinkan analisis data tabular berukuran kecil secara efisien [9][10]. Dengan memanfaatkan model *pretrained* seperti TabNet dan TabPFN, analisis prediktif dapat dilakukan terhadap dataset e-IPO yang terbatas namun dinamis. Pendekatan ini tidak hanya meningkatkan akurasi prediksi performa saham pasca-IPO, tetapi juga menghasilkan *insights* yang relevan bagi investor, regulator, pelaku pasar, serta perusahaan yang berencana melakukan IPO.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

1. Faktor-faktor apa saja yang paling berpengaruh terhadap performa saham e-IPO dalam jangka pendek hingga panjang?
2. Bagaimana model *pretrained* berbasis *transformer* seperti TabNet dan TabPFN dapat digunakan untuk memprediksi performa saham e-IPO di Indonesia pada berbagai jangka waktu (1 hari, 1 minggu, 1 bulan, 6 bulan, dan 1 tahun)?
3. Bagaimana efektivitas penggunaan model TabPFN pada dataset e-IPO yang kecil dan dinamis dibandingkan dengan TabNet dan model konvensional?

1.3 Tujuan Penelitian

Penelitian ini memiliki tujuan sebagai berikut:

1. Mengidentifikasi variabel-variabel penting yang memengaruhi performa saham e-IPO di bursa, baik dari sisi keuangan maupun non-keuangan.
2. Menerapkan model TabPFN dan TabNet untuk memprediksi kinerja saham e-IPO di Indonesia dalam lima rentang waktu yang berbeda.
3. Membandingkan kinerja prediksi TabPFN dengan TabNet dan model tradisional lainnya pada dataset e-IPO berukuran kecil dan heterogen.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Bagi akademisi: Menambah literatur dalam penerapan *artificial intelligence* berbasis *transformer* dalam prediksi pasar modal, khususnya untuk dataset kecil dan terbatas.
2. Bagi pelaku pasar: Memberikan wawasan akan faktor-faktor yang memengaruhi kinerja saham e-IPO, serta model AI yang dapat membantu memutuskan partisipasi pada e-IPO.
3. Bagi perusahaan yang ingin IPO: Menyediakan *insight* berbasis data terkait faktor-faktor yang perlu diperhatikan agar dapat memaksimalkan kinerja saham setelah *listing*.

2 Kajian Teori

2.1 Attention Mechanism

Attention mechanism adalah komponen kunci dalam arsitektur *deep learning* yang memungkinkan model untuk menimbang bagian-bagian input yang paling relevan dalam proses pengambilan keputusan. Secara umum, *attention* bekerja dengan mengkalkulasi skor relevansi antara suatu query Q dan sekumpulan keys K , kemudian menggunakannya untuk mengambil representasi dari values V . Dalam konteks data tabular, *attention* dimodifikasi agar bisa memilih fitur-fitur yang relevan pada

setiap observasi.

2.2 TabNet

TabNet adalah arsitektur *deep learning* berbasis *decision step* dan *attention mechanism* yang dirancang khusus untuk menangani data tabular. Model ini diperkenalkan oleh Arik dan Pfister (2019) dan menawarkan kemampuan seleksi fitur yang *sparsity-aware* serta interpretabilitas secara *end-to-end*, menandakan TabNet hanya memilih beberapa fitur yang dapat dipelajari di mask $M[i] \in \mathbb{R}^{B \times D}$ [11]. i menunjukkan langkah ke- i dalam *encoder*, jumlah elemen matriks ini berarti 1, yang berarti *importance* sebuah fitur dievaluasi secara relatif.

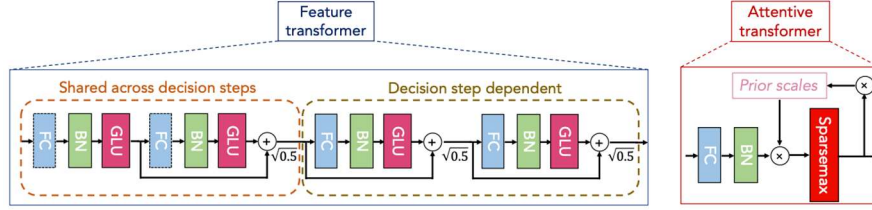


Figure 2: Arsitektur TabNet

2.2.1 Feature Transformer

Feature Transformer adalah *Multi-Layer Perceptron* (MLP) dalam arsitektur TabNet yang bertanggung jawab untuk menghasilkan representasi nonlinier dari fitur input. MLP ini terdiri atas *Fully Connected* (FC) layer, *Batch Normalization* (BN), dan fungsi aktivasi *Gated Linear Unit* (GLU). Secara matematis, input vektor fitur awal $x \in \mathbb{R}^d$ diproses melalui formula feature transformer (f_i), $[d[i], a[i]] = f_i(M[i], f)$. Digunakan *Ghost Batch Normalization*, dengan tujuan mempertahankan rata-rata varian yang rendah di antara *mini-batch*. Dalam implementasinya, ukuran batch dilambangkan sebagai B_v , dan momentum yang digunakan normalisasi ini dinyatakan sebagai m_B . Selain itu, sebelum dilakukan penggabungan pada residual connection, diterapkan normalisasi skalar sebesar $\sqrt{0.5}$ untuk stabilitas *learning process* [12].

Feature Transformer sendiri terdapat dua jenis *hidden features*, yaitu $a[i]$ untuk prior scale di *Attentive Transformer* dan keluaran keputusan $d[i]$, dihasilkan dari keseluruhan proses *Feature Transformer*, dan merupakan hasil akhir dari representasi yang digunakan untuk prediksi.

2.2.2 Attentive Transformer

Feature Transformer digunakan untuk menghasilkan *masking* berdasarkan fitur output dari langkah sebelumnya. Nilai *prior scale* $P[i - 1]$ dengan formula $P[i] = \prod_{j=1}^i (\gamma - M[j])$ merepresentasikan seberapa sering suatu fitur telah dipakai di langkah sebelumnya dan akan mempengaruhi pembuatan *mask*. Semakin besar nilai *prior* pada suatu fitur, semakin kecil kemungkinan fitur tersebut akan dipilih kembali. Parameter $P[i - 1]$ dapat disesuaikan menggunakan parameter γ , yang berfungsi untuk mengontrol fleksibilitas pemilihan fitur [13]. Terkhusus untuk langkah pertama (step ke-0), prior scale $P[0]$ diinisialisasi sebagai matriks satuan $1^{B \times D}$, yang berarti semua fitur dianggap belum digunakan.

Sebagai tahap akhir, hasil dari *Attentive Transformer* disempurnakan menggunakan fungsi normalisasi *sparsemax*, $L_{sparse} = \sum_{i=1}^{Nsteps} \sum_{b=1}^B \sum_{j=1}^D \frac{-M_{b,j}[i] \log(M_{b,j}[i] + \epsilon)}{Nsteps \cdot B}$. Berbeda dengan *softmax*, *sparsemax* menghasilkan distribusi yang jarang (sparse), yaitu hanya memberikan nilai positif pada sebagian kecil fitur, sementara sisanya bernilai nol. Pendekatan ini bertujuan untuk mengontrol tingkat *sparsity* dalam proses seleksi fitur, sekaligus memperkuat fokus model terhadap kolom-kolom yang memang memiliki pengaruh kuat terhadap target.

2.3 TabFPN

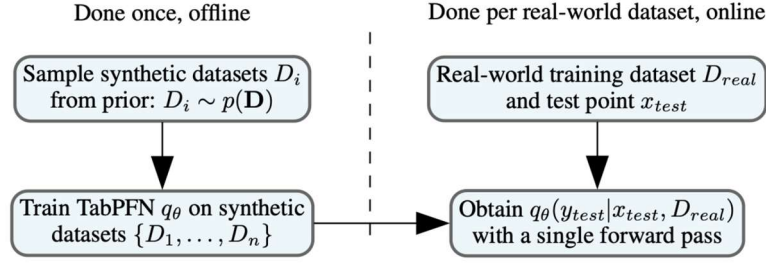


Figure 3: Arsitektur TabFPN

TabFPN (*Tabular Prior-Data Few-Shot Neural Network*) adalah metode berbasis transformer yang mampu melakukan prediksi supervised (klasifikasi dan regresi) pada data tabular dalam sekali *forward pass*, dengan mengeluarkan probabilitas prediksi sebagai berikut $\hat{p}(y_{test}|x_{test}, D) = \text{TabPFN}(x_{test}, D)$ [14]. Ini dikarenakan model sudah dilatih secara *meta-learning* dengan menggunakan dataset dari prior berupa *Structural Causal Models* (SCM), $D_i \sim p(D)$. Setiap dataset TabPFN akan dilatih untuk meminimalkan kesalahan prediksi, yaitu $\mathcal{L} = -\sum_{(x_{test}, y_{test})} \log \hat{p}(y_{test}|x_{test}, D_{train})$.

Tujuan dari pre-training ini adalah untuk mendekati inferensi Bayesian atas prior generatif tersebut. Secara teori, prediksi ideal yang ingin dicapai adalah $P(y_{test}|x_{test}, D) = \int P(y_{test}|x_{test}, D, \theta)p(\theta|D)d\theta$. Namun karena komputasi integral ini secara eksplisit tidak praktis (terutama dalam kasus *high-dimensional parameter space*), maka TabPFN belajar mengaproksimasi hasil akhir dari integral tersebut dengan cara langsung memetakan input dataset dan titik uji ke distribusi prediktif yang relevan melalui jaringan transformer.

3 Solusi Usulan

Dalam studi ini, pendekatan berbasis *attention mechanism* untuk data tabular, yaitu TabNet (Arik & Pfister, 2019) dan TabPFN (Hüser et al., 2023), dipilih sebagai metode utama untuk membangun model prediksi kinerja saham e-IPO. Dataset yang digunakan berisi 223 observasi dan lebih dari 10 fitur, yang merepresentasikan konteks data berukuran kecil namun bersifat multivariat, dengan karakteristik umum data pasar sekunder IPO yang bersifat volatil dan tidak linier. Keduanya menggunakan struktur *neural network* berbasis representasi selektif atas fitur input, namun memiliki karakteristik arsitektural, proses pelatihan, serta asumsi inferensial yang berbeda secara fundamental.

Sebagai pembandingan, studi ini juga menyertakan algoritma *machine learning* konvensional berbasis *decision tree*, yaitu *Extreme Gradient Boosting* (XGBoost), untuk mengevaluasi sejauh mana pendekatan berbasis *attention* mampu memberikan kinerja prediktif yang superior dalam konteks data tabular berukuran kecil.

4 Metodologi

4.1 Dataset

Teknik pengumpulan data dalam penelitian ini dilakukan melalui proses penyalinan informasi dan teknik *web scraping* dari situs e-ipo.co.id yang merupakan platform utama dalam penyampaian informasi terkait Penawaran Umum Perdana (*Initial Public Offering*) di Indonesia. Dataset yang dikumpulkan terdiri dari 223 entri, dimana setiap entri mewakili satu perusahaan. Fitur-fitur yang dikumpulkan meliputi informasi seperti sektor perusahaan, papan pencatatan, struktur kepemilikan, rasio keuangan (seperti ROE, leverage, equity), harga saham saat IPO, serta performa saham pada

periode 1 hari, 1 minggu, 1 bulan, 6 bulan, dan 1 tahun setelah IPO. Selain itu, informasi terkait underwriter, frekuensi keterlibatannya, kualitas, dan keberhasilannya dalam menangani IPO juga disertakan.

Dataset tersebut kemudian diproses melalui tahap pembersihan dan pengkodean untuk mempersiapkan pelatihan model klasifikasi. Target klasifikasi adalah apakah performa saham suatu perusahaan setelah IPO tergolong “positif” atau “negatif”. Penilaian ini tidak hanya didasarkan pada return satu tahun setelah IPO, tetapi juga mempertimbangkan perubahan harga saham pada beberapa horizon waktu, yaitu saat IPO, 1 hari, 1 minggu, 1 bulan, dan 6 bulan setelah IPO. Khusus untuk horizon 1 hari, pendekatan yang digunakan adalah klasifikasi multi-kelas dengan kelas tambahan yaitu ARA (Auto Reject Atas). Dengan demikian, label klasifikasi disusun berdasarkan tren keseluruhan performa saham dalam kurun waktu tersebut. Contoh distribusi target klasifikasi dapat dilihat pada Tabel 2.

Table 1: Contoh Data

Nama Perusahaan	Papan Pencatatan	Total Aset	Pendapatan	Kinerja (1Y)
PT Geoprima Solusi Tbk	Pengembangan	65.683,6 juta rupiah	1.949,42 juta rupiah	-32
PT Pertamina Geothermal Energy Tbk	Utama	37,1 triliun rupiah	4,36 triliun rupiah	40

Table 2: Distribusi Target Klasifikasi Berdasarkan Return 1 Tahun

Kelas	Jumlah	Persentase
Tidak Sukses	149	66.82%
Sukses	74	33.18%

Untuk proses pelatihan dan pengujian model, dataset utama dibagi menjadi data *training* dan *testing*. Selanjutnya, metode *K-fold Cross Validation* diterapkan pada tahap evaluasi untuk memastikan penilaian kinerja model yang robust dan *reliable* sebelum menghasilkan evaluasi akhir.

4.2 Metrik

Salah satu tantangan utama dalam dataset ini adalah adanya ketidakseimbangan distribusi kelas target. Fenomena ini terjadi secara konsisten di semua horizon waktu prediksi yang dianalisis (1 hari, 1 minggu, 1 bulan, 6 bulan, dan 1 tahun). Contoh ini dapat terlihat pada Tabel 2 terkait distribusi target klasifikasi pada 1 tahun.

Kondisi ketidakseimbangan yang persisten ini membuat metrik akurasi menjadi kurang ideal karena dapat memberikan gambaran performa yang terlalu optimis, meskipun model gagal mengidentifikasi kelas minoritas. Oleh karena itu, penelitian ini menetapkan F1-Score (Weighted Average) sebagai metrik evaluasi utama. F1-Score merupakan rata-rata harmonik dari *Precision* dan *Recall* sehingga mampu memberikan penilaian yang lebih seimbang terhadap kemampuan model.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 - \text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Dengan menggunakan *F1-Score* rata-rata tertimbang (*weighted average*), evaluasi kinerja tidak

hanya adil tetapi juga merefleksikan kemampuan model secara keseluruhan dengan memberi bobot pada setiap *F1-score* kelas sesuai dengan jumlah sampelnya. Pendekatan ini memastikan perbandingan performa antar model di setiap horizon waktu menjadi lebih valid.

4.3 Alur Eksperimen

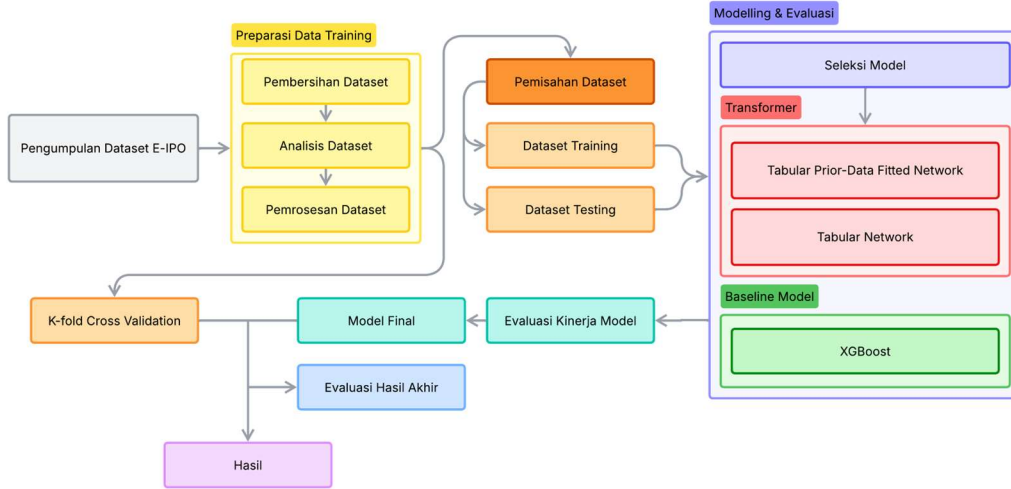


Figure 4: Diagram Alur Eksperimen

Alur eksperimen yang dilakukan oleh peneliti dapat dilihat pada Gambar 4. Tahapan Penelitian ini diawali dengan proses akuisisi data dari situs resmi e-IPO (<https://e-ipo.co.id/id/ipo/index>), di mana data terkait perusahaan yang mengajukan IPO disalin ke dalam format tabular dan kemudian dibersihkan untuk keperluan analisis. Setelah data siap digunakan, dilakukan pemisahan dataset menjadi data latih (*training*) dan data uji (*testing*).

Setelah data latih digunakan untuk tahap *modelling* dan evaluasi. Pada tahap ini, tim peneliti melakukan eksperimen dengan tiga model yang berbeda. Pertama, model *machine learning* konvensional XGBoost digunakan sebagai *baseline*. Kemudian, dua model berbasis *transformer* yang dirancang khusus untuk data tabular, yaitu Tabular Prior-Data Fitted Network (TabPFN) dan Tabular Network (TabNet), dilatih dan dievaluasi untuk dibandingkan dengan *baseline*.

Setelah evaluasi awal, metode *K-fold Cross Validation* diterapkan untuk mendapatkan validasi kinerja yang lebih robust dan memilih model final. Proses eksperimen ini diakhiri dengan pemilihan model final berdasarkan evaluasi kinerja yang diperoleh dari proses *K-fold Cross Validation*, dimana hasil validasi tersebut sekaligus menjadi temuan dan hasil akhir dari penelitian ini.

4.4 Pengaturan Eksperimen

Pengaturan parameter untuk setiap model yang digunakan dalam eksperimen ini dirinci pada Tabel 3. Setiap model dikonfigurasi secara spesifik untuk mencapai performa optimalnya pada dataset e-IPO.

Table 3: Pengaturan Model Dalam Eksperimen

Model	Parameter	Nilai
-------	-----------	-------

XGBoost	tree_method predictor use_label_encoder eval_metric	gpu_hist gpu_predictor False logloss
TabPFN	-	Tidak memerlukan parameter training (model bersifat <i>meta-trained</i>).
TabNet	n_d (lebar layer keputusan) n_a (lebar layer attention) n_steps (jumlah layer keputusan) gamma (koefisien relaksasi fitur) lambda_sparse (koefisien sparsity) optimizer scheduler max_epochs batch_size	16 16 5 1,5 0,0001 Adam (0,02) ReduceLROnPlateau (patience=5, factor=0.5) 25 16

Untuk model TabNet, pelatihan final dilakukan hingga epoch ke-25, yang merupakan epoch terbaik yang ditemukan selama tahap tuning untuk menghindari overfitting. Sementara itu, untuk XGBoost, pengaturan `tree_method` dan `predictor` dioptimalkan untuk memanfaatkan akselerasi GPU jika tersedia, guna mempercepat proses komputasi.

Perlu dicatat bahwa beberapa parameter, terutama `batch_size` dan `epoch`, disesuaikan secara dinamis untuk setiap horizon waktu prediksi (label) yang berbeda. Penyesuaian ini dilakukan untuk mencari konfigurasi paling optimal guna memaksimalkan kinerja model pada masing-masing tugas klasifikasi yang spesifik.

5 Pembahasan

5.1 Exploratory Data Analysis (EDA)

5.1.1 Pengaruh Tipe *Underwriter* dengan Kinerja Harga Saham e-IPO

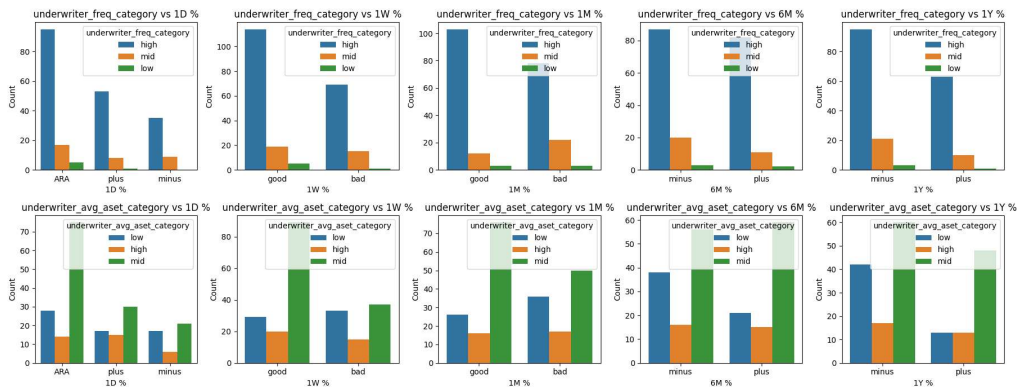


Figure 5: Perbandingan Frekuensi dan Rata-Rata Aset *Underwriter* Berdasarkan Performa

Underwriter dengan frekuensi tinggi memiliki volume aktivitas yang besar namun tidak secara konsisten berkorelasi positif dengan kinerja *return* jangka pendek IPO. Sedangkan, *underwriter* dengan frekuensi dan aset rendah lebih sering menangani saham yang berkinerja buruk.

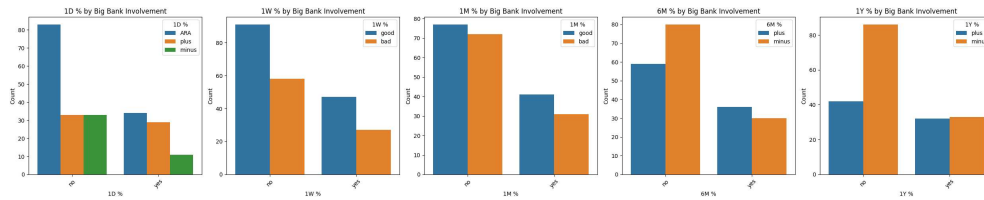


Figure 6: Bar Chart Keterlibatan Big Bank terhadap Kinerja Saham

Big bank adalah underwriter/perusahaan sekuritas milik lembaga keuangan ternama, seperti PT Mandiri Sekuritas. Di kasus ini, saham dengan *big bank* secara jangka panjang menunjukkan distribusi performa yang lebih seimbang antara *plus* dan *minus*. Sedangkan untuk *short-term*, saham *non-big bank* memiliki jumlah Auto Reject Atas (ARA) lebih banyak, menandakan IPO *non-big bank* cenderung lebih spekulatif dan agresif. Ini dikarenakan *big banks* lebih baik dalam melakukan valuasi perusahaan sehingga kinerja saham IPO yang dibawakan lebih stabil.

5.1.2 Pengaruh Jumlah Aset dengan Kinerja Harga Saham e-IPO

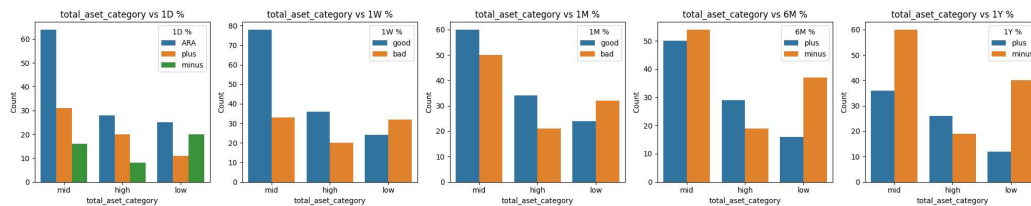


Figure 7: Bar Chart Kategori Jumlah Aset terhadap Kinerja Saham

Perusahaan dengan total aset yang tinggi cenderung memiliki performa yang lebih baik secara keseluruhan. Sementara, perusahaan dengan aset rendah menunjukkan kinerja saham yang lebih buruk, terutama dalam jangka menengah hingga panjang.

5.1.3 Faktor yang Mempengaruhi Emiten Beraset Rendah

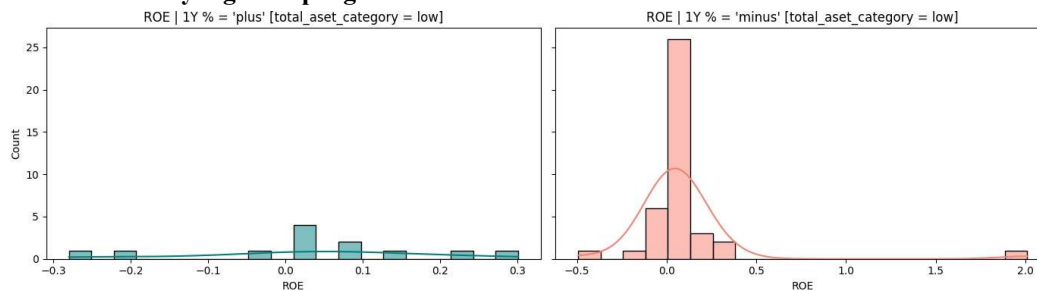


Figure 8: Distribusi ROE Berdasarkan Kinerja Saham 1 Tahun pada Emiten dengan Aset Rendah

Di antara emiten dengan total aset yang rendah, mayoritas saham yang mengalami penurunan kinerja satu tahun pasca-IPO menunjukkan nilai *Return on Equity* (ROE) yang rendah atau negatif. Hal ini menunjukkan bahwa profitabilitas yang rendah adalah salah satu faktor utama terhadap buruknya kinerja saham pada kelompok emiten beraset kecil.

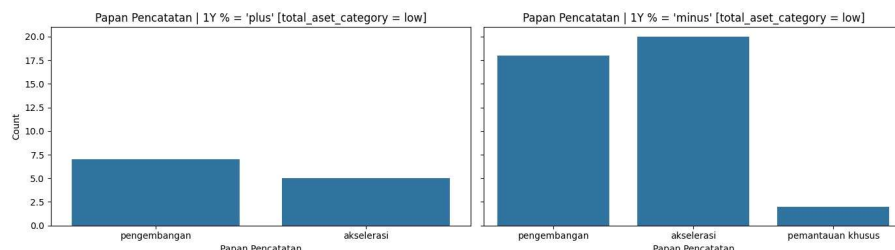


Figure 9: Distribusi Papan Pencatatan Berdasarkan Kinerja Saham pada Emiten Aset Rendah

Seluruh emiten beraset rendah yang tercatat di papan “*Pemantauan Khusus*” mengalami penurunan harga saham dalam satu tahun pasca-IPO. Oleh karena itu, keberadaan emiten di papan merupakan salah satu indikator kuat terhadap tingginya risiko kinerja perusahaan yang negatif secara jangka panjang

5.1.4 Beberapa Variabel yang Tidak Terlihat Berkorelasi

Dalam proses *Exploratory Data Analysis (EDA)*, beberapa variabel tambahan telah diuji untuk melihat kemungkinan korelasinya terhadap performa saham e-IPO. Namun, hasil analisis menunjukkan bahwa variabel-variabel tersebut tidak memberikan hubungan yang signifikan atau konsisten. Di antaranya adalah kuartal saat perusahaan melakukan IPO, indeks volatilitas pasar (VIX index), rencana penggunaan dana hasil IPO, serta beberapa metrik keuangan seperti NPM, PER, dan PBV. Hal ini mengindikasikan bahwa faktor-faktor fundamental atau makroekonomi tertentu mungkin tidak sepenuhnya menentukan kinerja awal saham di pasar, dan bahwa sentimen investor, reputasi underwriter, serta kondisi penawaran dan permintaan saat IPO justru bisa memiliki peran yang lebih dominan.

5.2 Analisis Hasil Eksperimen dan Pengujian

5.2.1 Perbandingan Kinerja Model

Pada tahap awal, penelitian ini menetapkan model machine learning konvensional, XGBoost, sebagai baseline untuk perbandingan. Selanjutnya untuk meningkatkan performa prediksi, dilakukan eksperimen menggunakan model-model berbasis transformer yang dirancang untuk data tabular, yaitu TabPFN dan TabNet.

Kinerja model dievaluasi pada berbagai horizon waktu menggunakan metrik F1-score rata-rata tertimbang (*weighted average*), yang memberikan gambaran performa model secara keseluruhan mempertimbangkan jumlah data pada tiap kelas. Penting untuk dicatat bahwa prediksi untuk horizon 1 hari (1D) merupakan klasifikasi multi-kelas (3 kelas), sedangkan horizon waktu lainnya (1 minggu hingga 1 tahun) adalah klasifikasi biner (2 kelas). Hasil Perbandingan F1-Score disajikan dalam Tabel 4.

Table 4: Perbandingan *F1-Score* Model (Weighted Average)

Model	F1-Score (1 Hari)	F1-Score (1 Minggu)	F1-Score (1 Bulan)	F1-Score (6 Bulan)	F1-Score (1 Tahun)
XGBoost	0,41	0,59	0,57	0,56	0,63
TabPFN	0,40	0,63	0,66	0,63	0,72
TabNet	0,59	0,65	0,65	0,63	0,73

Hasil pada Tabel 4 menunjukkan bahwa model-model berbasis transformer secara umum lebih unggul dibandingkan *baseline* XGBoost, terutama pada horizon waktu 1 minggu ke atas. TabNet secara konsisten melampaui XGBoost di semua horizon waktu. TabPFN, meskipun kinerjanya sebanding dengan XGBoost pada horizon 1 hari, menunjukkan keunggulan yang jelas di semua horizon lainnya. Hal ini mengonfirmasi efektivitas arsitektur *transformer* untuk menangani dataset e-IPO.

Saat membandingkan kedua model *transformer*, terlihat adanya perbedaan performa. TabNet menunjukkan keunggulan pada skenario klasifikasi multi-kelas (1 Hari) dengan skor 0,59, jangka pendek (1 Minggu) dengan skor 0,65, dan jangka panjang (1 Tahun) dengan skor 0,73. Sebaliknya,

TabPFN terbukti sedikit lebih unggul pada horizon waktu 1 bulan dengan skor 0,66. Pada horizon 6 bulan, performa keduanya tercatat identik dengan skor 0,63.

5.2.2 Analisis Pengaruh Jangka Waktu Prediksi

Analisis pada berbagai horizon waktu menunjukkan pergeseran faktor klasifikasi yang jelas, dari sentimen pasar jangka pendek ke fundamental perusahaan jangka panjang. Pada jangka pendek (1 hari – 1 minggu), dimana kinerja saham cenderung didorong oleh sentimen pasar, TabNet menunjukkan keunggulan yang jelas. Kemampuannya untuk memodelkan interaksi fitur yang kompleks kemungkinan besar lebih efektif dalam menangkap pola-pola non-linier yang menandakan adanya kondisi yang terjadi pada pasar.

Memasuki jangka menengah dan panjang (1 bulan – 1 tahun), fokus pasar bergeser ke arah evaluasi fundamental perusahaan. Pada tahap ini, TabPFN menjadi lebih kompetitif dan bahkan unggul pada horizon 1 bulan, sementara TabNet kembali memimpin pada horizon 1 tahun. Keberhasilan kedua model transformer pada periode ini menunjukkan kemampuan superior mereka dalam mengidentifikasi fitur-fitur penting yang berkaitan dengan data fundamental perusahaan, yang menjadi faktor penentu kinerja saham seiring berjalannya waktu.

5.2.3 Interpretasi Model

Untuk memahami faktor-faktor yang mendorong keputusan model, tim peneliti menggunakan *Local Interpretable Model-agnostic Explanations* (LIME). Analisis ini bertujuan untuk membongkar “kotak hitam” dan mengidentifikasi fitur apa saja yang dianggap penting oleh model saat membuat prediksi untuk data spesifik.

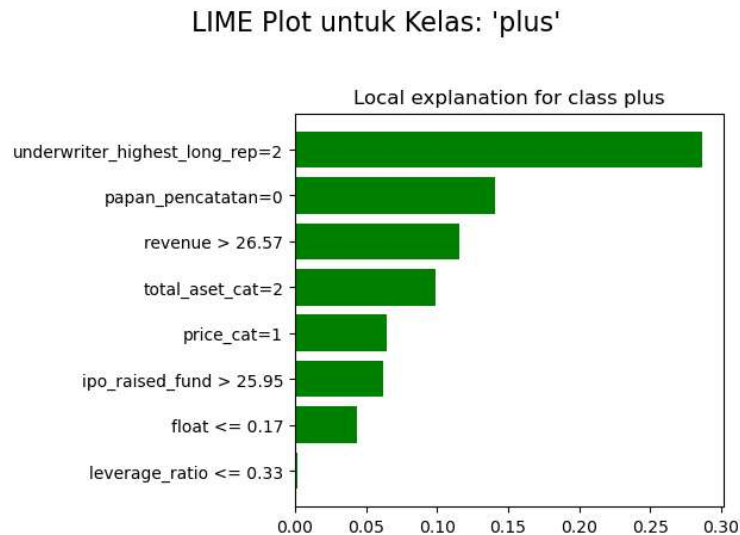


Figure 10: Visualisasi LIME untuk Prediksi Kinerja Positif (1 Tahun) pada Model TabNet

Analisis LIME selanjutnya memberikan wawasan mendalam mengenai faktor-faktor yang mendorong prediksi model untuk kelas “plus” (sukses) dan “minus” (gagal) pada kasus-kasus spesifik.

Untuk prediksi kelas “plus”, model secara signifikan mengandalkan fundamental perusahaan yang kuat. Faktor pendukung utama adalah `underwriter_highest_long_rep=2` yang menandakan perusahaan yang didukung oleh underwriter dengan rekam jejak jangka panjang terbaik. Faktor lainnya yang turut mendukung adalah `papan_pencatatan=0` (tercatat di papan utama), `revenue > 26,57` (pendapatan tinggi), dan `leverage_ratio <= 0,33` (rasio utang yang rendah). Hal ini

menunjukkan model telah belajar untuk mengasosiasikan keberhasilan IPO dengan kualitas fundamental dan reputasi penjamin emisi.

6 Penutup

6.1 Kesimpulan

Penelitian ini menunjukkan bahwa penggunaan model Tabular Transformer *dengan attention mechanism*, khususnya TabPFN dan TabNet, mampu memberikan hasil prediksi yang lebih akurat dalam menganalisis performa saham e-IPO di Indonesia dibandingkan dengan model konvensional seperti XGBoost. Keunggulan utama dari pendekatan ini terletak pada kemampuannya dalam secara selektif memfokuskan perhatian pada fitur-fitur penting melalui mekanisme *attention*, sehingga mampu menangkap pola-pola kompleks dalam data tabular yang bersifat multivariat dan memiliki ukuran kecil seperti dalam kasus e-IPO.

Model TabPFN, yang merupakan pendekatan probabilistik berbasis transformer, dan TabNet, yang menggunakan *sparse attention* untuk feature selection, terbukti mampu menghasilkan performa klasifikasi yang lebih stabil dan andal. Hasil ini memperkuat potensi penggunaan model *deep learning* pada domain keuangan, khususnya dalam konteks pasar modal Indonesia. Dengan demikian, pendekatan ini dapat menjadi sebuah solusi yang menjanjikan dalam mendukung pengambilan keputusan investasi di masa depan.

6.1 Saran

Saran untuk penelitian selanjutnya, bisa melakukan *benchmarking* dengan model transformer tabular lainnya. Beberapa model alternatif yang relevan untuk dibandingkan antara lain SAINT, FT-Transformer, dan *Transformer-Encoders*. *Benchmarking* ini penting untuk melihat bagaimana masing-masing model menangani tantangan pada data tabular dengan jumlah fitur terbatas namun kompleks, seperti pada kasus saham e-IPO.

Terakhir, tentunya perlu adanya edukasi masyarakat mengenai e-IPO dan risiko investasi di pasar modal. Meskipun model prediktif seperti TabPFN dan TabNet mampu memberikan rekomendasi berbasis data, masyarakat tetap memerlukan pemahaman dasar mengenai cara kerja pasar, potensi keuntungan, dan risiko kerugian. Oleh karena itu, hasil penelitian ini dapat dijadikan sebagai dasar untuk mengembangkan platform atau dashboard edukatif yang menjelaskan insight model dengan cara yang mudah dimengerti oleh publik non-teknis.

Referensi

- [1] Bursa Efek Indonesia. (n.d.). Beranda. <https://www.idx.co.id/id> (Diakses pada 10 Juli 2025)
- [2] Kustodian Sentral Efek Indonesia. (n.d.). Beranda. <https://www.ksei.co.id/> (Diakses pada 10 Juli 2025)
- [3] Rahmanto, M. R., Suroso, A. I., & Bandono, B. (2024). *The Successful Implementation of the Electronic System of Initial Public Offering ("E-IPO") in the Indonesian Capital Market*. *Indonesian Journal of Business and Entrepreneurship*, 10(1), 85. <https://doi.org/10.17358/ijbe.10.1.85>
- [4] Otoritas Jasa Keuangan. (n.d.). Beranda. <https://www.ojk.go.id/id/Default.aspx> (Diakses pada 10 Juli 2025)
- [5] PT Bursa Efek Indonesia. (n.d.). *Overview of e-IPO system. e-IPO Indonesia*. Retrieved July 10, 2025, from <https://www.e-ipo.co.id/docs/OVERVIEW.html>
- [6] Ariyanto, N. R., Haryetti, & Fathoni, A. F. (2020). *Detection Herding Behaviour on post initial public offering: An empirical study on Sharia stocks in Indonesia Stock Exchange [PDF]*. *South East Asia Journal of Contemporary Business, Economics and Law*, 21(5), 273–277. Retrieved July 10, 2025, from https://seajbel.com/wp-content/uploads/2020/07/SEAJBEL21_272.pdf
- [7] Singh, S. (2022, May 24). *Fear of missing out (FOMO) in IPO investments*. LinkedIn. <https://www.linkedin.com/pulse/fear-missing-out-fomo-ipo-investments-swati-singh-cdadf/>
- [8] Fuadi, A. A., & Pasaribu, P. (2024). *IPO long-term performance: Evidence from Indonesia Stock Exchange 2019–2020*. *Equity: Jurnal Keuangan dan Investasi Islam*, 27(2), 106–123. <https://doi.org/10.34209/equ.v27i2.9382>
- [9] Arik, S. Ö., & Pfister, T. (2019). *TabNet: Attentive interpretable tabular learning [Preprint]*. arXiv.

<https://arxiv.org/abs/1908.07442>

[10] Hüser, M., Schäfer, C., Wilming, J., & Macke, J. H. (2022). *TabPFN: A transformer that solves small tabular classification problems in a second [Preprint]*. arXiv. <https://arxiv.org/abs/2207.01848>

[11] Arik, S. Ö., & Pfister, T. (2021). *TabNet: Attentive interpretable tabular learning*. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 8, pp. 6679–6687). AAAI Press. <https://aaai.org/papers/06679-tabnet-attentive-interpretable-tabular-learning/>

[12] Gorishniy, Y., Rubachev, I., Khrulkov, V., & Babenko, A. (2021). *Revisiting deep learning models for tabular data*. Advances in Neural Information Processing Systems, 34, 18932–18943. https://proceedings.neurips.cc/paper_files/paper/2021/file/f15e177e418fcd3c3290e5b5d86c1ce-Paper.pdf

[13] Somepalli, G., Goldblum, M., Goldstein, T., & Sharma, A. (2021). *Improved neural networks for tabular data via row attention and contrastive pre-training*. arXiv preprint arXiv:2106.01342. <https://doi.org/10.48550/arXiv.2106.01342>

[14] Hüser, M., Schäfer, C., Wilming, J., & Macke, J. H. (2023). *TabPFN: A transformer that solves small tabular classification problems in a second*. Proceedings of the 40th International Conference on Machine Learning (ICML 2023), PMLR 202, 17783–17809. <https://proceedings.mlr.press/v202/huser23a.html>

Lampiran

Table 5: Detail Fitur Dataset

Fitur	Deskripsi
Ticker	Simbol unik saham perusahaan di Bursa Efek Indonesia
Nama Perusahaan	Nama lengkap perusahaan yang melakukan IPO
Tanggal Listing	Tanggal pertama kali saham perusahaan diperdagangkan di bursa (tanggal IPO)
Sektor	Kategori industri atau sektor ekonomi tempat perusahaan beroperasi
Underwriter	Perusahaan penjamin emisi yang membantu proses IPO perusahaan
Papan Pencatatan	Jenis papan di BEI tempat saham dicatatkan (Utama, Pengembangan, Akselerasi)
Ownership	Informasi kepemilikan saham sebelum IPO, bisa berupa struktur pemilik mayoritas atau institusi tertentu
Listed Share	Jumlah total saham yang dicatatkan di bursa setelah IPO (dalam unit lembar saham)
Persen Saham Ditawarkan	Persentase dari total saham yang dijual ke publik saat IPO
Rencana Penggunaan Dana IPO	Rencana perusahaan dalam menggunakan dana yang diperoleh dari IPO
Valuation Currency	Mata uang yang digunakan untuk valuasi keuangan (biasanya IDR atau USD).
Total Aset	Jumlah total aset perusahaan saat IPO

Total Liabilitas	Total kewajiban/hutang perusahaan pada saat IPO
Pendapatan	Total pendapatan perusahaan (revenue) dalam periode keuangan terakhir sebelum IPO
Laba Kotor	Pendapatan dikurangi dengan biaya pokok penjualan
Laba Berjalan (Bukan Komprehensif)	Laba bersih yang diperoleh perusahaan dalam laporan laba rugi
Harga Final	Harga final saham yang ditetapkan untuk IPO
1D %	Persentase perubahan harga saham satu hari setelah listing
1W %	Persentase perubahan harga saham satu minggu setelah IPO
1M %	Persentase perubahan harga saham satu bulan setelah IPO
6M %	Persentase perubahan harga saham enam bulan setelah IPO
1Y %	Persentase perubahan harga saham satu tahun setelah IPO
