

SCRAPING WEBISTE MARKETPLACE

**Makalah ini disusun dalam rangka memenuhi Tugas Mata Kuliah
“Kecerdasan Artificial”**



Kelapa Muda

Hadron Herman Cong	(221111825)
Christian Tiovanto	(221111153)
Devon Marvellous Loen	(221110226)
Jerry	(221110417)
Jerry Wijaya	(221110969)

UNIVERSITAS MIKROSKIL
Jl. M.H Thamrin No.140, Pusat Ps., Kec. Medan Kota, Kota Medan,
Sumatera Utara 20212
TA 2023/2024

Daftar isi

Bab 1 Pendahuluan.....	3
1.1 Latar Belakang	3
1.2 Tujuan	4
Bab 2 Pembahasan.....	5
2.1 Pembahasan Deskripsi Dan Progress.....	5
2.2 Pembahasan Pembagian Tugas	6
Bab 3 Lampiran.....	8
3.1 Code Program.....	8
3.2 UI (User Interfaces)	9
Bab 4 Penutup	10
4.1 Kesimpulan.....	10
4.2 Saran.....	10

Bab 1

Pendahuluan

1.1 Latar Belakang

Dalam proyek ini, kami mengembangkan sebuah program menggunakan bahasa Python yang bertujuan untuk melakukan Web Scraping pada beberapa situs web marketplace terkemuka di Indonesia, yaitu Tokopedia dan Bukalapak. Tujuan utama dari program ini adalah untuk mengumpulkan informasi produk dari situs-situs tersebut secara otomatis, dan kemudian menyimpannya dalam format file Excel untuk kemudian dianalisis atau digunakan untuk keperluan lainnya.

Metode yang digunakan untuk melakukan Web Scraping adalah dengan menggunakan pustaka atau Library Python yang telah tersedia, seperti Selenium untuk mengendalikan browser web dan BeautifulSoup untuk melakukan Parsing pada dokumen HTML dari halaman web yang diakses.

Pertama-tama, program akan meminta masukan pengguna berupa nama produk yang ingin dicari. Selanjutnya, pengguna akan diberikan pilihan untuk menerapkan filter tertentu, seperti produk terbaru, termurah, atau dengan peringkat tertinggi. Setelah filter dipilih, program akan membangun URL sesuai dengan pilihan pengguna dan melakukan akses ke halaman web marketplace yang sesuai.

Kemudian, program akan mengekstrak informasi yang relevan dari halaman web, seperti nama produk, harga, peringkat, jumlah penjualan, lokasi toko, dan nama toko. Informasi ini akan diolah dan disimpan dalam bentuk struktur data yang sesuai, seperti daftar atau kamus.

Setelah selesai melakukan scraping dari situs web, program akan mengeksport data yang telah dikumpulkan ke dalam format file Excel. Hal ini memungkinkan pengguna untuk dengan mudah melihat dan menganalisis informasi produk yang telah diambil dari berbagai situs web marketplace.

Selain program berbasis Python untuk melakukan Web Scraping, proyek ini juga melibatkan pengembangan antarmuka pengguna (UI) berbasis mobile menggunakan Flutter. Flutter dipilih karena kemampuannya untuk membangun aplikasi mobile yang responsif dan menarik dengan satu basis kode yang dapat dijalankan pada perangkat Android dan iOS.

Aplikasi mobile ini akan berfungsi sebagai platform bagi pengguna untuk memasukkan nama produk yang ingin dicari serta memilih filter yang diinginkan secara lebih mudah dan intuitif. Dengan antarmuka pengguna yang user-friendly, aplikasi ini akan mempermudah pengguna dalam mengakses dan memanfaatkan data yang telah dikumpulkan oleh program web scraping.

Dengan demikian, proyek ini memberikan cara efisien dan otomatis untuk mengumpulkan informasi produk dari beberapa situs web marketplace terkemuka di Indonesia, serta menyediakan fleksibilitas untuk menganalisis dan menggunakan data tersebut sesuai kebutuhan pengguna. Penambahan antarmuka pengguna berbasis mobile dengan Flutter akan semakin mempermudah dan meningkatkan pengalaman pengguna dalam menggunakan program ini.

1.2 Tujuan

Adapun tujuan yang ingin dicapai pada makalah ini sebagai berikut :

1. **Mengumpulkan Data Produk** : Tujuan utama dari proyek ini adalah untuk mengumpulkan informasi lengkap mengenai produk-produk yang tersedia di berbagai situs web marketplace seperti Tokopedia dan Bukalapak. Informasi yang dikumpulkan mencakup nama produk, harga, peringkat, jumlah penjualan, lokasi toko, dan nama toko.
2. **Menfasilitasi Analisi** : Dengan mengumpulkan data produk dari beberapa situs web marketplace, program ini memungkinkan pengguna untuk melakukan analisis komprehensif terkait dengan produk-produk yang ada di pasar. Analisis tersebut dapat mencakup perbandingan harga, popularitas produk, atau tren penjualan dari waktu ke waktu.
3. **Mendukung Pengambilan Keputusan** : Informasi yang diperoleh dari proses **Web Scraping** dapat membantu pengguna dalam mengambil keputusan yang lebih baik terkait dengan pembelian produk. Misalnya, pengguna dapat membandingkan harga produk dari beberapa toko atau mengetahui penilaian dan ulasan dari pembeli sebelumnya,
4. **Mengotomatiskan Proses** : Program ini mengotomatiskan proses pengumpulan informasi produk dari situs web marketplace, menghemat waktu dan usaha yang diperlukan untuk mencari dan mencatat informasi tersebut secara manual.
5. **Membantu Riset Pasar** : Data yang dikumpulkan dapat digunakan untuk riset pasar yang lebih mendalam. Pengguna dapat mengidentifikasi tren pasar, preferensi konsumen, atau kesenjangan pasar yang dapat dimanfaatkan untuk mengembangkan strategi pemasaran atau produk baru.
6. **Menghasilkan Laporan** : Program ini memungkinkan pengguna untuk membuat laporan yang rapi dan terstruktur mengenai informasi produk. Laporan ini dapat digunakan untuk keperluan internal, presentasi, atau berbagi dengan pihak lain yang terkait.

Dengan Demikian, Tujuan Utama dari proyek ini adalah untuk menyediakan alat yang efektif bagi pengguna untuk mengumpulkan, menganalisis dan menggunakan informasi produk dari berbagai situs web marketplace, dengan tujuan akhir untuk mendukung pengambilan keputusan yang lebih baik dan mengembangkan pemahaman yang lebih dalam tentang pasar.

Bab 2

Pembahasan

2.1 Deskripsi Dan Progress

Topik utama dalam proyek ini adalah melakukan **web scraping** data dari tiga marketplace ternama: **Tokopedia** dan **Bukalapak**. Kami berhasil mengumpulkan informasi lengkap mengenai setiap produk, termasuk nama produk, harga, rating, nama toko, dan jumlah penjualan. Selain itu, kami juga melaksanakan pengembangan fitur filter untuk memudahkan pencarian berdasarkan harga terendah dan tertinggi, memastikan pengguna dapat menyesuaikan pencarian mereka sesuai kebutuhan. Pada tahap terakhir, kami menyajikan data yang terkumpul dengan rapi dan mudah dimengerti.

Langkah 1: Perencanaan

- Identifikasi sumber data yang akan di-scrape, seperti situs web Tokopedia dan Bukalapak.
- -Buat daftar informasi yang akan diambil dari setiap situs web, misalnya nama produk, harga, rating, nama toko, dan jumlah penjualan.
- Tentukan teknologi yang akan digunakan, seperti library Python BeautifulSoup atau Selenium.
- Menyusun rencana untuk mengembangkan antarmuka pengguna (UI) mobile menggunakan Flutter.

Langkah 2: Pengembangan Kode

- Buat skrip Python untuk melakukan web scraping dari masing-masing situs web.
- Implementasikan fitur filter untuk mencari harga terendah dan tertinggi.
- Uji kode secara berkala untuk memastikan keakuratan dan keandalan dalam mengambil data.

Langkah 3: Integrasi dan Analisis Data

- Gabungkan data yang di-scrape dari semua situs web ke dalam satu struktur data yang seragam.
- Lakukan analisis data untuk mengidentifikasi tren pasar atau informasi penting lainnya

Langkah 4: Pengembangan UI Mobile dengan Flutter

- Menyusun desain antarmuka pengguna (UI) mobile menggunakan Flutter, termasuk halaman pencarian produk, filter, dan hasil pencarian.
- Mengimplementasikan fitur-fitur seperti input pencarian, filter harga, dan tampilan hasil pencarian.
- Melakukan pengujian untuk memastikan antarmuka berfungsi dengan baik dan sesuai dengan kebutuhan pengguna.

Dengan mengikuti langkah-langkah ini, proyek web scraping dapat berkembang secara bertahap dan mencapai tujuan yang diinginkan dengan lebih efektif.

2.2 Pembagian Tugas

Dalam pengembangan proyek web scraping untuk mengumpulkan data produk dari situs web marketplace dan pembuatan UI (Mobile Flutter), pembagian tugas menjadi hal krusial untuk memastikan efisiensi dan kualitas hasil akhir. Berikut adalah penjelasan tentang bagaimana tugas-tugas dibagi di antara anggota tim:

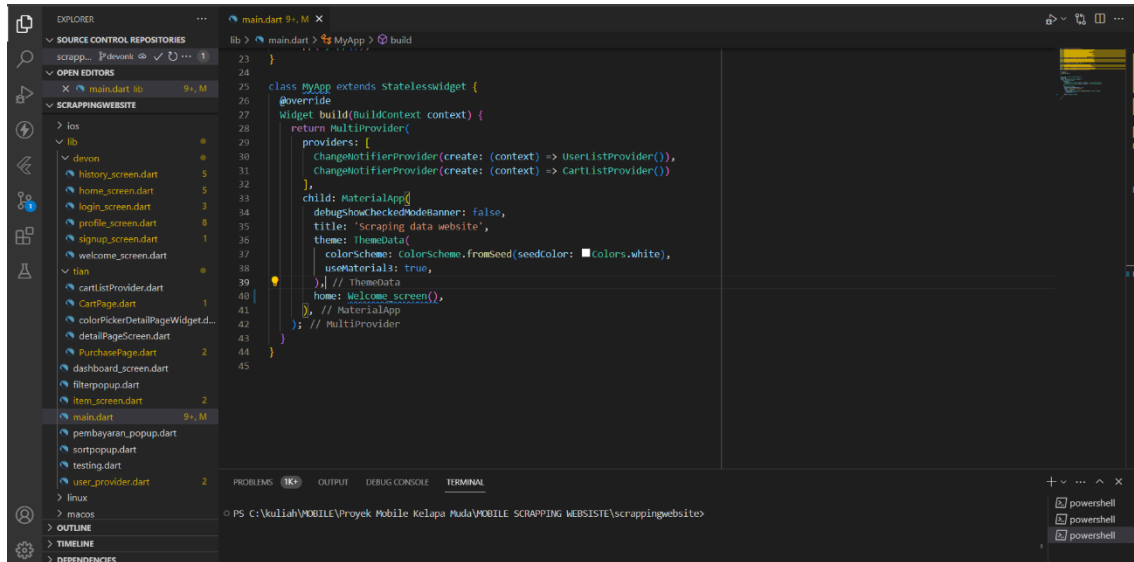
Nama	Tugas	Deskripsi	Estimasi (days)	Prioritas
Devon Marvellous Loen (221110226)	Web Scraping TokoPedia	Bertanggung jawab untuk mengembangkan Code Program Web Scraping Filter Harga Termurah/Termahal pada situs web Tokopedia dan mengembangkan Code Program pada situs web Tokopedia	4	HIGH
Jerry (221110417)	Web Scraping TokoPedia dan Bukalapak	Bertanggung jawab untuk mengembangkan Code Program Web Scraping Pada situs web Tokopedia dan Bukalapak serta Menggabungkan Code Program dari Ketiga Website tersebut	4	HIGH
Hadron Herman Cong (221111825)	Web Scraping Bukalapak	Bertanggung jawab untuk mengembangkan Code Program Web Scraping pada situs web Bukalapak serta Pengembangan Filter Harga Termurah/Termahal pada website tersebut	4	HIGH
Christian Tiovanto (221111153)	Web Scraping TokoPedia	Bertanggung jawab untuk mengembangkan Code Program Web Scraping pada situs web Tokopedia .	4	HIGH
Jerry Wijaya (221110969)	Web Scraping Bukalapak	Bertanggung jawab untuk mengembangkan Code Program Web Scraping pada situs web Bukalapak .	4	HIGH

Nama	Tugas	Deskripsi	Estimasi (days)	Prioritas
Devon Marvellous Loen (221110226)	Flutter Mobile Front End and Impement Back End	Saya bertanggung jawab mengembangkan antarmuka pengguna (UI) mobile menggunakan Flutter dan mengintegrasikan backend untuk web scraping dari Tokopedia, dan Bukalapak. Tugas utama saya mencakup perancangan UI responsif, implementasi fitur pencarian produk dan filter harga, serta memastikan integrasi yang efektif antara frontend dan backend untuk pengelolaan data scraping yang jelas dan informatif dalam aplikasi Flutter.	15	HIGH
Jerry (221110417)	Backend implementation on scraping data and user structure	Saya bertanggung jawab untuk membuat api backend yang akan digunakan oleh front end untuk mendapatkan data dari hasil scrapingan yang ia mau cari. Scrapingan data dilakukan di dalam aplikasi backend. Membuat user untuk bisa mendaftar dan login, mengembalikan jwt token yang sebagai pertanda bahwa itu adalah user dari aplikasi yang kita buat. Membuat fitur keranjang dan pembelian untuk hasil data yang di scraping.	15	HIGH
Hadron Herman Cong (221111825)	Design figma for Flutter Mobile Front End and scraping data detail Bukalapak	Saya bertanggung jawab untuk membuat design antar muka (UI) front end pada figma dan menscraping data detail pada Bukalapak	15	HIGH
Christian Tiovento (221111153)	Flutter Mobile Front End and Impement Back End	Saya bertanggung jawab untuk membuat UI dengan menggunakan Flutter dan mengintegrasikan UI tersebut dengan Aplikasi BackEnd yang sudah dibuat	15	HIGH
Jerry Wijaya (221110969)	Design figma for Flutter Mobile Front End and scraping data detail Bukalapak	Saya bertanggung jawab untuk membuat design antar muka (UI) front end pada figma dan menscraping data detail pada Bukalapak	15	HIGH

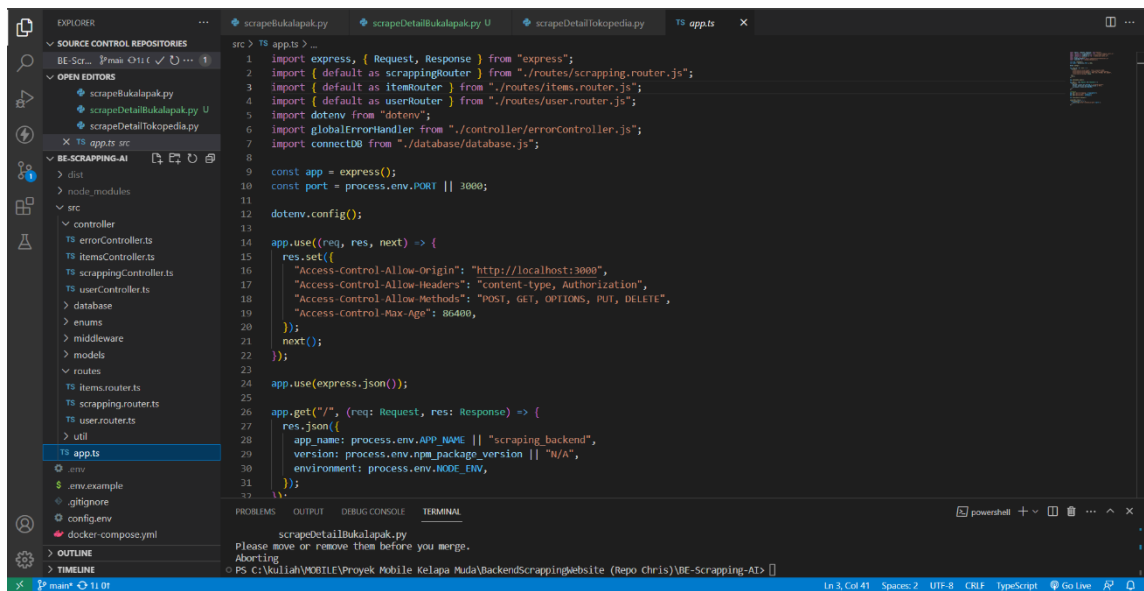
Bab 3

Lampiran

3.1 Code Program

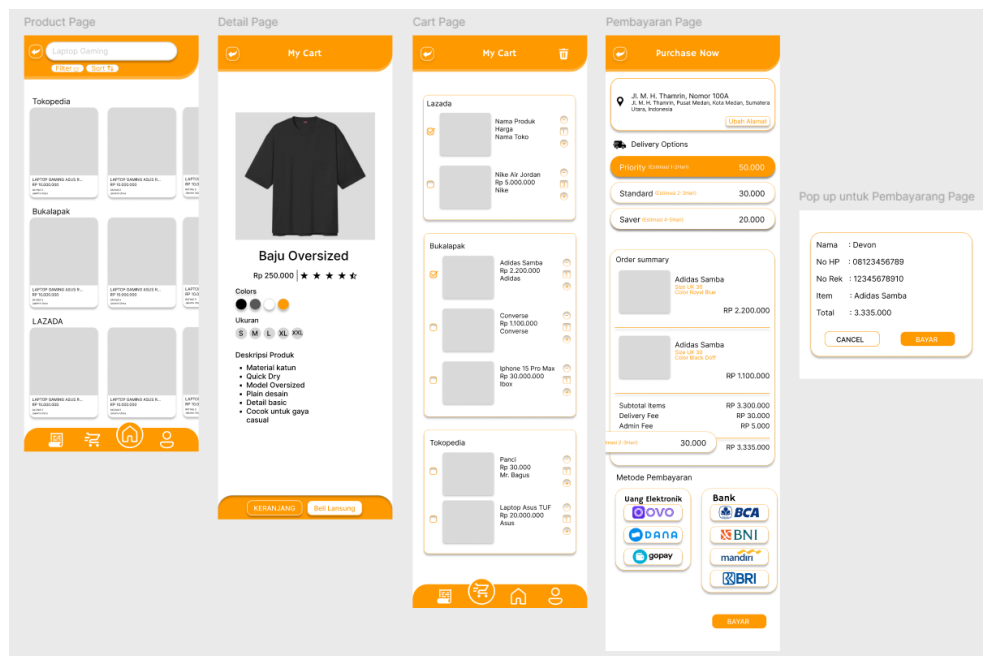
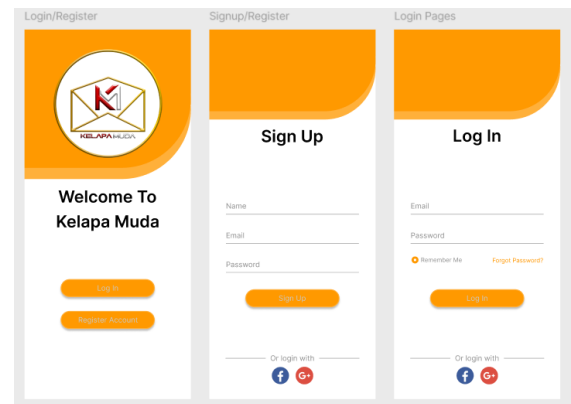
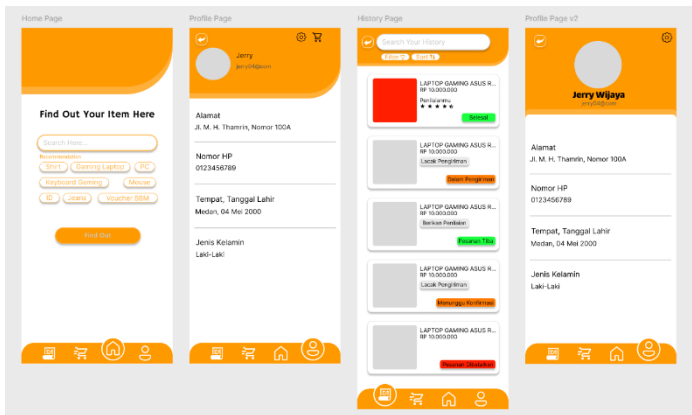


```
23 }
24
25 class MyApp extends StatelessWidget {
26   @override
27   Widget build(BuildContext context) {
28     return MultiProvider(
29       providers: [
30         ChangeNotifierProvider(create: (context) => UserlistProvider()),
31         ChangeNotifierProvider(create: (context) => CartlistProvider()),
32       ],
33       child: MaterialApp(
34         debugShowCheckedModeBanner: false,
35         title: 'Scraping data website',
36         theme: ThemeData(
37           colorScheme: ColorScheme.fromSeed(seedColor: Colors.white),
38           useMaterial3: true,
39         ), // ThemeData
40         home: Welcome screen(),
41       ), // MaterialApp
42     ); // MultiProvider
43   }
44 }
45
```



```
src > ts app.ts > ...
1 import express, { Request, Response } from "express";
2 import { default as scrapingRouter } from "./routes/scraping.router.js";
3 import { default as itemRouter } from "./routes/items.router.js";
4 import { default as userRouter } from "./routes/user.router.js";
5 import dotenv from "dotenv";
6 import errorHandler from "./controller/errorController.js";
7 import connectDB from "./database/database.js";
8
9 const app = express();
10 const port = process.env.PORT || 3000;
11
12 dotenv.config();
13
14 app.use((req, res, next) => {
15   res.set({
16     "Access-Control-Allow-Origin": "http://localhost:3000",
17     "Access-Control-Allow-Headers": "Content-type, Authorization",
18     "Access-Control-Allow-Methods": "POST, GET, OPTIONS, PUT, DELETE",
19     "Access-Control-Max-Age": 86400,
20   });
21   next();
22 });
23
24 app.use(express.json());
25
26 app.get("/", (req: Request, res: Response) => {
27   res.json({
28     app_name: process.env.APP_NAME || "scraping_backend",
29     version: process.env.npm_package_version || "1.0.0",
30     environment: process.env.NODE_ENV,
31   });
32 });
33
```


3.1 UI (User Interfaces)



LINK GITHUB BERISI VIDIO PENJELASAN DAN CODE PROGRAM

Bab 4

Penutup

4.1 Kesimpulan

Dari hasil pembahasan di atas, dapat disimpulkan bahwa:

1. **Manfaat Komunikasi Terbuka dengan Metode Web Scraping:** Proses pengembangan web scraping tidak hanya memfasilitasi pengambilan data yang akurat dari marketplace seperti Tokopedia dan Bukalapak, tetapi juga mendorong komunikasi terbuka dan intensif di antara anggota tim. Pengembangan antarmuka pengguna (UI) mobile menggunakan Flutter memberikan platform yang memungkinkan anggota tim untuk berkolaborasi secara langsung, menyampaikan ide, dan mengintegrasikan masukan dengan cepat, yang esensial untuk meningkatkan kualitas dan ketepatan solusi akhir.
2. **Pemecahan Masalah dan Penyesuaian Solusi:** Dengan pendekatan iteratif dalam pengembangan UI mobile dan implementasi backend untuk web scraping, tim dapat dengan fleksibel mengevaluasi dan menyesuaikan solusi mereka sesuai dengan perubahan yang terjadi. Ini memungkinkan tim untuk merespons dengan cepat terhadap tantangan teknis maupun kebutuhan pengguna, memastikan bahwa aplikasi yang dikembangkan tidak hanya efektif dalam pengambilan dan pengelolaan data, tetapi juga memberikan informasi yang relevan dan bermanfaat bagi pengguna akhir.

4.2 Saran

Adapun saran yang dapat diberikan oleh penulis terkait dengan pembahasan di atas adalah:

1. **Web Scraping dalam Tim :** *Web scraping* dapat dilakukan dengan sukses oleh tim apabila semua anggota tim memahami pentingnya kerjasama, keaktifan, dan kedisiplinan dalam mengumpulkan dan mengelola data dari berbagai sumber online. Komunikasi yang terus-menerus dan terjadwal antara anggota tim akan memperkuat kohesi tim dan memastikan arah yang jelas dalam pencapaian tujuan *web scraping*. Oleh karena itu, penting bagi tim untuk menetapkan strategi yang jelas, mengidentifikasi sumber data yang relevan, dan merencanakan proses scraping secara terstruktur.
2. **Pengembangan Fitur Tambahan :** Selain fitur pencarian dan filter harga, Nantinya disarankan untuk menambahkan fitur tambahan seperti notifikasi harga, perbandingan produk, dan integrasi dengan alat analisis untuk memberikan nilai tambah bagi pengguna dan meningkatkan utilitas aplikasi.
3. **Pemantauan dan Pembaruan Berkala :** Menerapkan sistem pemantauan untuk melacak performa aplikasi dan mengidentifikasi masalah secara proaktif. Disarankan untuk merencanakan pembaruan berkala yang mencakup perbaikan bug, peningkatan fitur, dan penyesuaian terhadap perubahan di marketplace target.