

**Analisis Pemanfaatan *Predictive Analytic* Untuk Pemilihan Target Investasi
di Perusahaan-Perusahaan *Startup* Menggunakan Algoritma *Logistic
Regression* dan *K-Means***

TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar Magister
dari Program Studi PJJ S2 MANAJEMEN

Disusun oleh:

IDA BAGUS PUTU ARIARTHA

2403221047



**Universitas
Telkom**

**PJJ S2 MANAJEMEN
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS TELKOM
BANDUNG
2025**

JUDUL PENELITIAN

*Analysis of Predictive Analytics utilization for Selecting Investment Targets in
Startup Companies Using Logistic Regression and K-Means Algorithms*

Diajukan sebagai salah satu syarat untuk memperoleh gelar Magister
dari Program Studi PJJ S2 Manajemen

Disusun oleh:

IDA BAGUS PUTU ARIARTHA

2403221047



**Universitas
Telkom**

**PJJ S2 MANAJEMEN
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS TELKOM
BANDUNG
2025**

HALAMAN PENGESAHAN

**Analisis Pemanfaatan *Predictive Analytic* Untuk Pemilihan Target Investasi
di Perusahaan-Perusahaan *Startup* Menggunakan Algoritma *Logistic
Regression* dan *K-Means***

Diajukan sebagai salah satu syarat untuk memperoleh gelar Magister
dari Program Studi PJJ S2 Manajemen

Disusun oleh:

IDA BAGUS PUTU ARIARTHA

2403221047



Pembimbing

Dr. DADAN RAHADIAN, S.T., M.M.

**PJJ S2 MANAJEMEN
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS TELKOM
BANDUNG
2025**

HALAMAN PERNYATAAN

Dengan ini saya, Ida Bagus Putu Ariartha, NIM 2403221047, menyatakan bahwa tesis dengan judul “**Analisis Pemanfaatan *Predictive Analytic* Untuk Pemilihan Target Investasi di Perusahaan-Perusahaan *Startup* Menggunakan Algoritma *Logistic Regression* dan *K-Means***” adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan kecuali melalui pengutipan sesuai dengan etika keilmuan yang berlaku. Saya bersedia menanggung risiko/sanksi yang dijatuhkan kepada saya apabila ditemukan pelanggaran terhadap etika keilmuan dalam tesis saya ini.

Bandung, 15 Januari 2025

Yang membuat pernyataan,



Ida Bagus Putu Ariartha

NIM 2403221047

KATA PENGANTAR

Puji syukur penulis panjatkan kehadapan Ida Sang Hyang Widhi Wasa / Tuhan Yang Maha Esa karena atas limpahan Rahmat-Nya penulis dapat menyelesaikan tesis dengan judul “**Analisis Pemanfaatan *Predictive Analytic* Untuk Pemilihan Target Investasi di Perusahaan-Perusahaan *Startup* Menggunakan Algoritma *Logistic Regression* dan *K-Means***”. Tujuan penulisan tesis ini yaitu untuk memenuhi salah satu syarat kelulusan jenjang pendidikan S-2 Program Studi PJJ S2 Manajemen Fakultas Ekonomi dan Bisnis, Universitas Telkom Bandung.

Dalam penelitian ini, penulis banyak mendapatkan bimbingan, kritik, saran, dan motivasi yang sangat besar dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Dadan Rahadian selaku dosen pembimbing.
2. Ibu Irni Yunita dan Ibu Farida Titik Kristanti yang sudah banyak memberikan feedback untuk perbaikan proposal ini.
3. Ibu Maria Apsari Sugiat selaku Dosen Wali
4. Orang Tua, istri dan anak-anak yang sudah banyak memberikan dukungan selama penulis mengikuti perkuliahan.
5. Pihak-pihak lain yang ikut berkontribusi pada penelitian ini

ABSTRAK

Tujuan yang ingin dicapai dalam penelitian ini adalah mencari target *startup* yang bisa diberikan pendanaan, berdasarkan variabel yang berpengaruh terhadap kesuksesan sebuah *startup* menggunakan data *startup* yang ada di Crunchbase. Metode yang akan digunakan dalam penelitian ini adalah metode penelitian kuantitatif dengan pendekatan deskriptif, dimana penulis akan melakukan *predictive analysis* dengan menggunakan Crunchbase sebagai sumber data. Penelitian akan dilakukan dengan menguji apakah pendanaan yang diperoleh startup di seed stage berpengaruh terhadap kesuksesan sebuah *startup* (IPO, unicorn, diakuisisi, *late stage*)

Dalam penelitian ini, penulis menggunakan beberapa variabel data, diantaranya : *Company Name* sebagai variabel kunci, *variabel predictable* adalah status dari *startup* apakah sudah sukses/exit (IPO, unicorn, diakuisisi, *late stage*), variabel terkait karakteristik *startup*, variabel terkait aktifitas pendanaan. Populasi yang digunakan dalam penelitian ini adalah data *startup* yang ada dalam data Crunchbase pada akhir tahun 2023, yang berjumlah 3.000.000+ *startup*. Metode pemilihan sampel yang digunakan pada penelitian ini yaitu : *startup* yang terdaftar atau beroperasi di kawasan Asia Tenggara, *startup* yang bidang usahanya dibatasi pada beberapa sektor prioritas, *startup* yang didirikan pada periode tahun 2008-2023 dan *startup* yang mendapatkan pendanaan terakhir pada periode 2018-2023.

Dengan menggunakan *input variable* yang sudah didefinisikan, dan menggunakan algoritma *logistic regression* dapat diketahui variabel yang memiliki keterkaitan yang kuat terhadap status *startup* yang sukses. Setelah diketahui variabel yang berpengaruh terhadap kesuksesan *startup*, selanjutnya dilakukan *clustering* menggunakan algoritma *K-Means*. Dengan membuat *clustering* maka akan bisa diketahui kesamaan karakteristik *startup* yang diprediksi akan sukses di masa mendatang.

Kata kunci : *Venture capital, startup, logistic regression, K-Means, clustering*

ABSTRACT

Aim for this research is to find targeted startups to fund, based on startup's success variable using data on Crunchbase. This study uses quantitative research with descriptive approach, where the writer will use predictive analysis alongside Crunchbase as the data source. This study will test which variable affects a startup's success (IPO, acquired, unicorn, late stage)

On this study, the writer will use some variable data, some of it are: company's name as key variable, startup status (whether the startup already IPO, acquired, unicorn, late stage) as predictable variable, startup characteristic variable, and fund activity variable. 3.000.000+ startup data by the end of 2023 on Crunchbase will use as the population for this study. Sample sorting methods used in this study are: licensed startup or startup that operate in South East Asia, startups whose business fields are limited to several priority sectors, startup founded in 2008-2023, and startups that received their last funding in the period 2018-2023.

Using defined input variable and using logistic regression algorithm to find which variable has strong connection to a success startup. After success startup variable found then using clustering method and K-means algorithm, targeted fund startup will be determined. Using clustering will determine which startup have the same success profile in the future in one cluster.

Keywords : *Venture capital, startup, logistic regression, K-Means, clustering*

DAFTAR ISI

Halaman Pengesahan.....	iii
Halaman Pernyataan.....	iv
Kata Pengantar	v
Abstrak	vi
<i>Abstract</i>	vii
Daftar Isi.....	viii
Daftar Tabel.....	xi
Daftar Gambar.....	xii
BAB I PENDAHULUAN.....	1
1.1. Gambaran Umum Objek Penelitian.....	1
1.2. Latar Belakang Penelitian.....	4
1.3. Perumusan Masalah.....	10
1.4. Tujuan Penelitian.....	11
1.5. Manfaat Penelitian.....	11
1.5.1. Aspek Teoritis	11
1.5.2. Aspek Praktis.....	12
1.6. Sistematika Penelitian	12
BAB II TINJAUAN PUSTAKA	13
2.1 Teori Terkait Penelitian.....	13
2.1.1 Pengertian Investasi.....	13
2.1.2 Perusahaan <i>Startup</i>	13
2.1.3 Kriteria Sukses <i>Startup</i>	15
2.1.4 <i>Venture Capital</i>	16
2.1.5 Proses Pengambilan Keputusan di <i>Venture Capital</i>	19
2.1.6 Crunchbase.....	21
2.2 Penelitian Terdahulu.....	23
2.3 Kerangka Pemikiran	29
2.4 Hipotesis Penelitian	30
BAB III METODE PENELITIAN	32

3.1.	Jenis Penelitian	32
3.2.	Operasionalisasi Variabel.....	33
3.3.	Tahapan Penelitian	35
3.3.1.	<i>Business Understanding</i>	38
3.3.2.	<i>Data Understanding</i>	38
3.3.3.	<i>Data Preparation</i>	38
3.3.4.	<i>Modelling</i>	38
3.3.5.	<i>Evaluation & Development</i>	41
3.4.	Populasi dan Sampel	41
3.4.1.	Populasi	41
3.4.2.	Sampel	41
3.5.	Pengumpulan Data dan Sumber Data.....	42
3.6.	Teknik Analisis Data	43
Bab IV	Hasil Penelitian dan Pembahasan.....	46
4.1.	Deskripsi Data	46
4.2.	Profil Data Startup di Asia Tenggara	49
4.2.1.	Startup berdasarkan Negara Asal	49
4.2.2.	Startup berdasarkan tahun berdiri	50
4.2.3.	Startup berdasarkan sektor industri	50
4.2.4.	<i>Startup Exit</i>	51
4.2.5.	Startup Berhenti Operasi	52
4.2.6.	Startup Berdasarkan <i>Funding stage</i>	53
4.3.	Pengolahan Data.....	56
4.4.	Hasil Pengolahan Data	58
4.4.1.	Analisa <i>Scatter Plot</i>	58
4.4.2.	Analisa <i>Prediction Model</i>	59
4.4.3.	Hasil <i>Prediction Model</i>	62
4.4.4.	Hasil <i>Clustering</i>	64
4.5.	Pemanfaatan <i>Prediction Model</i> dan <i>Clustering</i>	69
Bab V	Kesimpulan dan Saran.....	70
5.1.	Kesimpulan.....	70

5.2. Saran.....	71
DAFTAR PUSTAKA.....	72

DAFTAR TABEL

Tabel 2.1 Tahapan <i>startup</i> dan pendanaan	18
Tabel 2.2 <i>Sourcing</i> target investasi.....	20
Tabel 2.3 Penelitian terdahulu	23
Tabel 3.1 Karakteristik Penelitian	33
Tabel 3.2 Variabel target	34
Tabel 3.3 Variabel karakteristik <i>startup</i>	34
Tabel 3.4 Variabel aktifitas pendanaan.....	34
Tabel 3.5 <i>Data Science Tasks and Examples</i>	40
Tabel 4.1 <i>Analytic Based Table</i>	56
Tabel 4.2 <i>Confusion matrix</i>	59
Tabel 4.3 <i>Confusion matrix data set training</i>	60
Tabel 4.4 Hasil <i>Test & Score</i>	60
Tabel 4.5 Hasil prediksi berdasarkan Negara	64
Tabel 4.6 Hasil prediksi berdasarkan sektor industri.....	64

DAFTAR GAMBAR

Gambar 1.1 Jumlah Pendanaan <i>Startup</i> di Indonesia	1
Gambar 1.2 <i>Fund</i> yang dikelola MDI Ventures	2
Gambar 1.3 <i>Portfolio</i> MDI Ventures	2
Gambar 1.4 <i>VC capital raised Southeast Asia</i>	4
Gambar 1.5 Penyebab kegagalan <i>startup</i>	5
Gambar 1.6 Prosentase kegagalan <i>startup</i> di Asia Tenggara	6
Gambar 1.7 Prosentase kegagalan <i>startup</i> di ASEAN berdasarkan <i>funding stage</i> ..	6
Gambar 1.8 Jumlah <i>startup unicorn</i> di Asia Tenggara.....	7
Gambar 1.9 Jumlah nilai <i>exit startup</i> teknologi di Asia Tenggara	7
Gambar 1.10 Jumlah <i>Startup</i> yang <i>exit</i>	8
Gambar 2.1 <i>Startup life cycle</i>	14
Gambar 2.2 Proses investasi VC	20
Gambar 2.3 Kerangka Pemikiran.....	30
Gambar 3.1 Proses Pemilihan Target Investasi	35
Gambar 3.2 <i>Data Science Process</i>	37
Gambar 3.3 Tahapan Pemodelan	38
Gambar 3.4 Jumlah Populasi	42
Gambar 3.5 Tahapan proses pengumpulan dan pengolahan data	43
Gambar 4.1 Data Crunchbase	46
Gambar 4.2 Struktur informasi profil perusahaan <i>startup</i>	47
Gambar 4.3 Struktur informasi pendanaan	47
Gambar 4.4 Data perusahaan <i>startup</i> di Crunchbase.....	48
Gambar 4.5 Data jumlah pendanaan di Crunchbase.....	48
Gambar 4.6 Proses <i>filtering</i> data.....	49
Gambar 4.7 <i>Startup</i> berdasarkan negara asal.....	49
Gambar 4.8 <i>Startup</i> berdasarkan tahun berdiri.....	50
Gambar 4.9 <i>Startup</i> berdasarkan sektor industri	51
Gambar 4.10 <i>Startup exit</i>	52
Gambar 4.11 Data <i>startup</i> berhenti beroperasi.....	53

Gambar 4.12 Data <i>startup</i> berdasarkan <i>funding stage</i>	54
Gambar 4.13 Periode waktu penerimaan pendanaan pertama	54
Gambar 4.14 Jumlah putaran pendanaan di tahap <i>seed</i>	55
Gambar 4.15 Total nilai pendanaan yang diperoleh <i>startup</i> berdasarkan tipe pendanaan terakhir.....	55
Gambar 4.16 Jumlah putaran pendanaan untuk <i>startup exit/closed</i>	56
Gambar 4.17 Variabel Input untuk Pemodelan	57
Gambar 4.18 Pemodelan dengan menggunakan Orange Data Mining.....	57
Gambar 4.19 Variabel yang digunakan dalam pemodelan	58
Gambar 4.20 <i>Scatter plot</i> waktu perolehan dan Jumlah putaran pendanaan fase <i>seed</i>	58
Gambar 4.21 <i>Scatter plot</i> waktu perolehan dan total pendanaan <i>fase seed</i>	59
Gambar 4.22 Kurva ROC	61
Gambar 4.23 Hasil <i>prediction model</i>	62
Gambar 4.24 Hasil prediksi berdasarkan waktu perolehan pendanaan	62
Gambar 4.25 Hasil prediksi berdasarkan putaran pendanaan.....	63
Gambar 4.26 Hasil prediksi berdasarkan total pendanaan.....	63
Gambar 4.27 Pemodelan <i>clustering startup</i>	65
Gambar 4.28 Jumlah <i>cluster K-means</i>	65
Gambar 4.29 Kluster berdasarkan waktu pendanaan dan jumlah putaran pendanaan	66
Gambar 4.30 Kluster berdasarkan umur <i>startup</i> dan waktu pendanaan.....	67
Gambar 4.31 Kluster berdasarkan umur <i>startup</i> dan jumlah putaran pendanaan..	68

BAB I

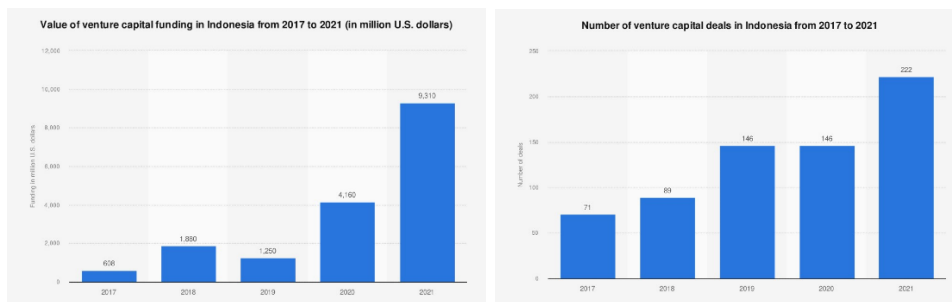
PENDAHULUAN

1.1. Gambaran Umum Objek Penelitian

Dalam beberapa tahun terakhir pertumbuhan *startup*, *venture capital (VC)* dan pendanaan di Indonesia mengalami peningkatan yang cukup signifikan. Berdasarkan data Statista (2022), ekosistem *startup* Indonesia menduduki peringkat kedua di kawasan Asia Tenggara setelah Singapura dalam hal jumlah *unicorn*. Indonesia memiliki lebih dari 2.300 *startup*, dengan lima *unicorn*.

Berdasarkan data Kementerian Kominfo Indonesia (2023), Pemerintah juga turut berperan aktif untuk mengembangkan *startup* lokal melalui Gerakan Nasional 1000 *Startup* Digital dan Platform HUB.ID, inisiatif Pemerintah ini bertujuan untuk membantu *startup* lokal yang masih berada dalam tahap awal untuk mengembangkan produk dan mencapai kesiapan pasar.

Pertumbuhan industri *startup* yang cukup pesat dalam beberapa tahun terakhir karena didukung oleh transformasi digital yang dilakukan di berbagai sektor. Berdasarkan data Statista (2022) Total pendanaan kepada *startup* di Indonesia pada tahun 2021 mencapai USD 9,3 miliar, yang dibagi menjadi 222 *deal*.



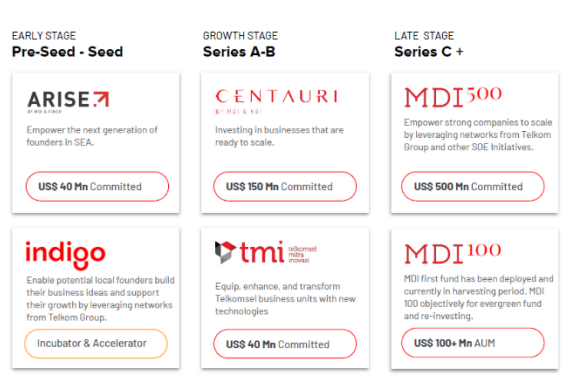
Gambar 1.1 Jumlah Pendanaan *Startup* di Indonesia

Sumber: Statista (2022)

Melihat perkembangan jumlah *startup* dan pendanaan yang diberikan, Telkom Indonesia pada tahun 2015 membentuk MDI Ventures, yang bertujuan untuk ikut memberikan kontribusi bagi pengembangan bisnis digital di Indonesia. *Venture capital* yang dibentuk ini memberikan pendanaan untuk *startup* di seluruh tahapan investasi untuk menjaga efisiensi dan memastikan tidak ada peluang yang terlewatkan, dan disamping itu MDI Ventures juga mengelola investasi dari investor pihak ketiga lainnya. *Unique selling point* yang ditawarkan oleh MDI adalah akses terhadap peluang bisnis baru melalui kolaborasi antara *startup* yang

diberikan pendanaan dan *investor* utama MDI yaitu Telkom, BUMN dan korporasi besar Indonesia lainnya (MDI Ventures, 2023).

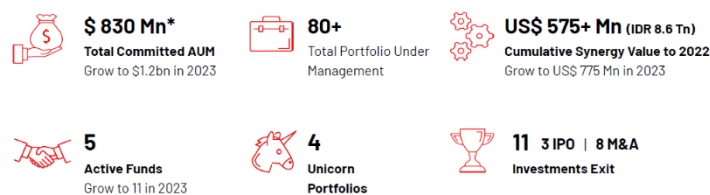
MDI Ventures saat ini mengelola beberapa *Fund*, dan juga diberikan mandat untuk mengelola program inkubator dan akselerator, dan juga *angel investor network*. MDI Ventures menyatukan *operation investment* menggunakan metode *one-team* dan *multiple-fund*, yang bertujuan untuk menjaga efisiensi dan tidak ada peluang investasi yang tertinggal.



Gambar 1.2 Fund yang dikelola MDI Ventures

Sumber : MDI Ventures (2023)

Sampai Triwulan 2 tahun 2023, MDI Ventures sudah memberikan pendanaan untuk kurang lebih 80+ *startup*, dengan total *committed asset under management* (AUM) mencapai 830 juta USD. Dari 80+ *startup* yang sudah diberikan pendanaan, tidak seluruhnya bisa memberikan *return* sesuai dengan yang diharapkan, saat ini baru ada 4 *startup* yang sudah masuk dalam kategori *unicorn*, dan 11 *startup* yang exit melalui IPO dan M&A (MDI Ventures, 2023).



Gambar 1.3 Portfolio MDI Ventures

Sumber : MDI Ventures (2023)

Dalam memberikan pendanaan kepada *startup*, MDI Ventures fokus pada beberapa sektor industri yang menjadi prioritas pengembangan bisnis digital Telkom Group. Sektor industri yang menjadi fokus diantaranya: *fintech*, *agriculture*, *logistics*, *big data & AI*,

education, media & entertainment, digital advertising, cloud computing, ecommerce, cyber security, dan IOT.

Berbeda dengan *venture capital* murni, yang hanya memiliki tujuan melakukan investasi hanya dari aspek keuntungan finansial, MDI Ventures sebagai *corporate venture capital* disamping memiliki tujuan keuntungan finansial, juga memiliki tujuan yang lebih strategis, dimana MDI Ventures berharap bisa mendapatkan akses teknologi dari *startup* yang diberikan pendanaan, dan mendapatkan keuntungan dari sinergi antara perusahaan induk dan *startup* yang diberikan pendanaan.

Adanya fenomena penurunan valuasi *startup* dan berkurangnya investor besar yang berminat untuk melakukan investasi di *startup* melalui *venture capital* dalam beberapa tahun terakhir memberikan dampak yang kurang bagus bagi MDI Ventures, diantaranya: imbal hasil yang tidak sesuai dengan yang diharapkan, ketidakpastian kapan *startup* akan *exit* dan juga tingginya *unrealized loss* yang disebabkan oleh penurunan valuasi *startup* yang mempengaruhi kinerja Perusahaan. Fenomena tersebut perlu diantisipasi, salah satunya dengan menerapkan kehati-hatian dalam proses pendanaan yang akan diberikan kepada *startup*.

MDI Ventures dalam mencari target *startup* yang akan diberikan pendanaan menggunakan beberapa tahapan, tahap pertama adalah *sourcing*, yaitu menyusun *long list startup* yang memenuhi kriteria yang sudah ditetapkan sebelumnya. Dari *long list startup* yang sudah disusun, selanjutnya dilakukan proses *pre-due diligence* (pre-DD), yaitu melakukan *scoring* dari setiap *startup* yang ada dalam *long list* dengan menggunakan beberapa parameter penilaian. Kemudian jika *startup* lolos dari tahap pre-DD selanjutnya akan dilakukan *assessment* yang lebih mendalam terhadap masing-masing *startup* (proses *due diligence*).

Pada tahapan *sourcing* untuk mencari kandidat *startup* yang akan menjadi target pendanaan, MDI Ventures menggunakan 2 metode yaitu: (a) *inbound*, dimana *startup* yang aktif menyampaikan proposal ke MDI Ventures, dan (b) *Outbound*, dimana MDI Ventures yang aktif mencari target *startup* untuk diberikan pendanaan melalui *networking* dengan VC, *investor* dan *event*. Dengan mengandalkan metode *inbound* dan *networking* untuk mendapatkan kandidat *startup*, ada kemungkinan MDI Ventures tidak mendapatkan kandidat *startup* yang terbaik yang memiliki potensi untuk bisa sukses di masa yang akan datang, karena *startup* yang ditawarkan mungkin terbatas, dibandingkan jumlah *startup* yang sudah lebih dari 3 juta.

Crunchbase merupakan *platform* yang menyediakan informasi mengenai perusahaan baik *private* maupun publik, yang memuat informasi mengenai *profile founder* dan *management team*, negara dimana perusahaan tersebut beroperasi, pendanaan yang diperoleh