POLITECNICO DI TORINO

Master's degree program

In Management Engineering

Master's thesis

AI on venture Capital

Supervisor

Prof. Riccardo Calcagno

Candidate

Achille Reale

A.A. 2024/2025

Indice

Introduction

Venture capital (VC) is one of the most dynamic and high-risk types of investment, playing a pivotal role in supporting the growth of startups and boosting the world's technological innovation. Different from traditional financing, VC focuses on early-stage and high-potential startups, often characterized by uncertainty and rapid scaling possibilities. Over the last few decades, this market has evolved, from an almost exclusively relationship-based activity to a more structured and data-driven one, driven by globalization, digitalization, and the increasing complexity of markets.

The traditional VC model has historically relied on the experience and intuition of general partners (GPs), who select startups based on personal networks, the so-called "gut feeling" and qualitative assessments of founders and market potential. However, as the number of startups continues to grow worldwide, and as investment opportunities become larger and more competitive, these methods alone are no longer sufficient to ensure sustained success. Investors face growing pressure to make faster and more accurate decisions while managing risk and maximizing returns for limited partners (LPs). In this context, technological innovation, particularly artificial intelligence (AI), has emerged as a transformative force capable of redefining how venture capital operates.

AI introduces the possibility of augmenting and, in some cases, partially replacing human judgment through data-driven insights, predictive modeling, and automation. Advanced machine learning algorithms can analyze massive amounts of unstructured data, including social media signals, patent filings, and market trends, to identify hidden opportunities or early signals of success. Natural language processing (NLP) models, meanwhile, enable automatic analysis of pitch decks, founder communications, and even public interviews to assess sentiment, consistency, and potential red flags. These technologies offer VCs new ways to prioritize deal flow, conduct due diligence, and monitor portfolio performance in real time.

Nevertheless, the integration of AI into venture capital is not without challenges. First, technical limitations such as data incompleteness, model opacity, and poor generalizability in new contexts highlight the importance of maintaining human oversight. Moreover, the presence of algorithmic bias and the difficulty in achieving true transparency raise ethical questions about fairness, diversity, and accountability in startup funding. If not properly addressed, these issues could undermine the legitimacy of AI-based decisions and erode trust among entrepreneurs and investors alike.

Beyond technical and ethical concerns, AI adoption is also reshaping the internal organization of VC firms. The traditional dominance of partners with strong networks and qualitative intuition is increasingly complemented, or even challenged, by data scientists and technical analysts. This evolution creates new skill requirements and hybrid roles, demanding that investment professionals learn to interpret and work alongside algorithmic outputs rather than rely solely on personal judgment. Furthermore, it introduces cultural tensions between more traditional investors and tech-driven teams, requiring firms to rethink their governance and internal dynamics.

On a broader level, AI has the potential to democratize access to capital. By enabling smaller investors to access sophisticated analysis tools and data previously available only to large funds, AI can foster a more diverse and competitive investment landscape. Emerging platforms and syndicate models allow a wider range of participants — including retail investors — to engage in early-stage investments, possibly reducing market concentration and opening opportunities to a more varied set of entrepreneurs worldwide. This democratization, however, must be balanced carefully with the need to maintain professional rigor and avoid the risks of speculative bubbles fueled by algorithmically driven herd behavior.

Looking to the future, the thesis explores different scenarios for the evolution of venture capital. Hybrid models, which combine algorithmic efficiency with human creativity and relationship-building, seem to offer the most balanced and resilient approach. While the idea of fully automated VC funds may appear attractive for their promise of speed and scalability, it remains limited by ethical, strategic, and trust-related constraints. The nuanced reality suggests that the greatest value lies in integration rather than the substitution of human expertise.

This work aims to analyze in depth how AI technologies can support and transform each stage of the VC process, from scouting and initial evaluation to due diligence, monitoring, and exit strategies. It will present a comprehensive review of AI methodologies, including machine learning, NLP, clustering techniques, and generative models, and discuss their concrete applications in the VC context. Several case studies of leading firms such as Tribe Capital, EQT Ventures, and GV (Google Ventures) illustrate how data-driven approaches are already shaping real-world investment practices.

Moreover, the thesis will address the limits and risks of AI in venture capital, proposing strategies to mitigate biases and maintain a healthy balance between automation and human oversight. By exploring future perspectives, including the potential for capital democratization and the emergence of hybrid models, this work hopes to provide a forward-looking framework for investors, founders, and policymakers interested in the evolving intersection of finance, technology, and entrepreneurship.

Ultimately, the objective of this thesis is to contribute to the ongoing dialogue on the future of venture investing, emphasizing that while AI represents an extraordinary tool for enhancing efficiency and insight, it should be viewed as an enabler rather than a replacement for human expertise. The combination of data-driven analysis with human intuition, relational capital, and strategic vision will likely define the next chapter of venture capital, opening new possibilities for innovation and economic growth on a global scale.

Chapter 1

Venture capital’s structure

Venture Capital is a form of private equity (PE) financing that provides resources and capital to startups with high potential. It is riskier than normal loans, but the returns can be extremely rewarding if the startup have success.

PE was born in the US after the Second World war and it grew rapidly on the 1970s and 1980s, especially for the ERISA Reform ("Prudent Man Rule" Clarification) which allowed pension funds to invest legally in riskier assets including VCs and PE. This led billions of dollars to flow in VC funds and so startups, scaling massively the industry, with Silicon Valley as the major hub.

Some of these startups such as Google, Microsoft, Facebook and so on now shape and lead the global innovation landscape and have capitalization of trillions of dollars.

1.1) Legal structure, stakeholders and investment logic

Most venture capital funds are structured as a limited partnership (LPs), as they provide a good balance between operational flexibility, tax efficiency and legal protection. This is ideal when managing high-risk, long-term investments in early-stage startups (ref. 1 - Metrick & Yasuda, *Venture Capital and the Finance of Innovation*, Journal of Economic Literature, 2010)

The two key stakeholders of the limited partnership are:

- General partner (GP): is the management position of the fund, it holds unlimited liability and is responsible for the operations such as investment decisions, portfolio management, and exits. In return it receives Management fees (around 2%) and Carried interest (around 20% of profits). As we will see later, GP is the main target that uses and benefits from the adoption of AI models.

- Limited partners (LPs): entity such as pension funds, endowment and foundations that contribute with large quantity of capital. LPs have no control over investment decisions but benefit from returns as defined by the partnership agreement. In return they receive basic return on investment, ROI (around 80% of profits).

The main ingredients for a good relationship between the two parties are trust, communication, transparency, alignment, professionalism and performance. Most of the time both parties exit the deal in a winning position, but there have been some cases that led to lawsuits and bankruptcy, worth mentioning are the Abraaj Group (2019), a major private equity/VC firm collapsed after misuse of LP funds, and Softbank's investments on WeWork, that turned out to be a historical bubble.

Startups are risky businesses to invest in; this is because most of them fail or underperform. The philosophy then is "high risk, high returns", meaning that even if most of the projects will not go through, the ones that do will produce incredible returns. In fact, one or two "unicorns" (startup valued > 1 B $) can cover the losses from the rest (ref. 2- Kerr, Nanda & Lerner, *Entrepreneurship as Experimentation*, Journal of Economic Perspectives, 2014).

Firstly (for the reasons mentioned above), VCs prioritize investments in startups with significant scalability potential and access to large or emerging markets, this is to ensure that successful venture can grow fast and capture substantial market share.

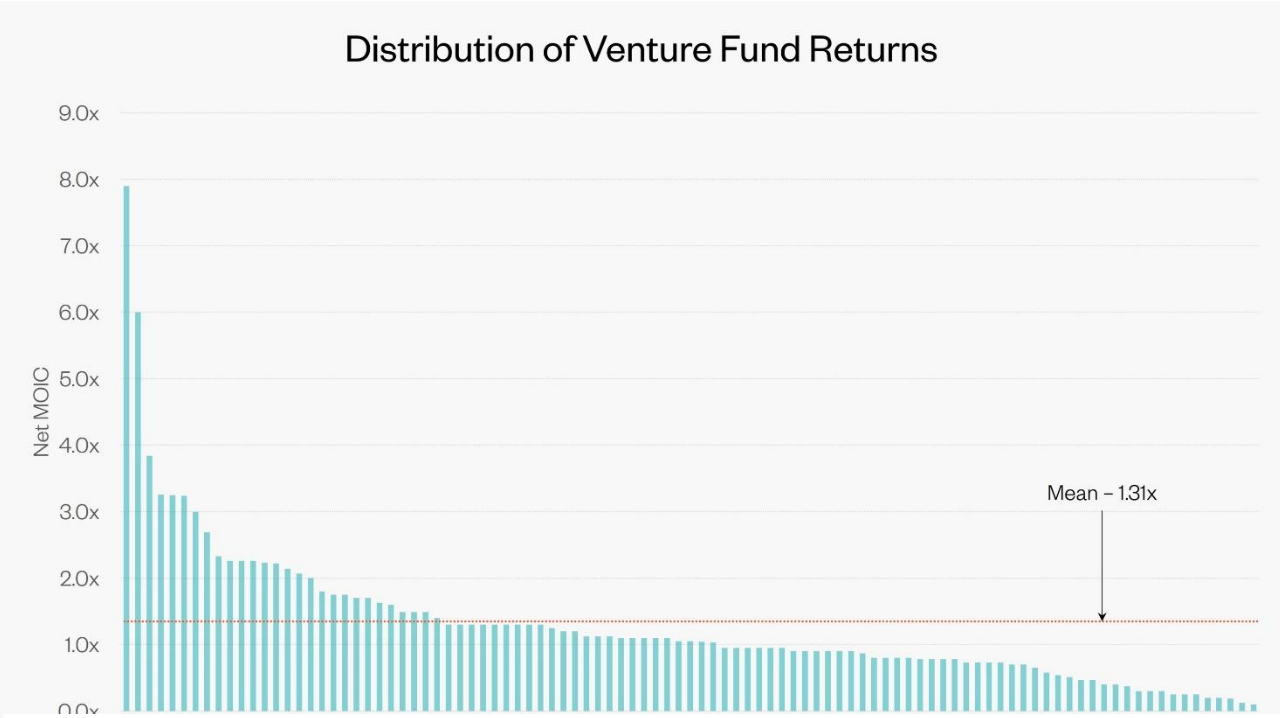


Figure 1

As a second critical factor VCs look for is the capabilities and characteristics of the startup's founding team, they must have strong execution and adaptability to uncertainty. A metaphor the incapsulate the concept is "bet on the jockey, not the horse".

Moving on, VCs' investments are based on stages, the earlier ones (pre-seed and seed) are characterized by a high risk and low amount of capital, while later stage rounds (round A and beyond) focus on scaling startups that have a product which is tested and ready for the market.

Furthermore, a clear and viable exit strategy is essential to the VC investment model. Since venture funds have a finite life (usually ten years) they need to generate returns via exits through acquisition, IPO, or secondary sale. The likelihood and timing of success significantly influence investment selection.

VCs also get involved actively in the companies in which they invest. Besides investing capital, they often serve on boards, offer strategic counsel, assist with recruitment, and leverage their networks to drive growth and future fund-raising. Active participation results in greater accountability. Interest alignment between entrepreneurs and VCs is facilitated by legal and financial tools including liquidation preferences, preferred stock, anti-dilution terms, and governance rights. The terms safeguard investors' capital and influence company decision-making arrangements.

In recent years, data analytics and artificial intelligence advancements have started to supplement traditional qualitative approaches, helping venture capitalists better detect potential startups, forecast market opportunities, and maximize portfolios.

To summarize, venture capital investment logic is a complex synergy of risk management, analysis of market opportunity, human capital evaluation, and financial structuring for the creation of superior long-term returns in an inherently uncertain environment

1.2) Investment phases

The investment in startups through stages is a crucial and structured process in which investors raise capital, allowing it to grow and reach its milestones. Stages are usually sequential, and each is characterized by distinct goals, expectations and investor profiles. Ranking up through stages, the startup not only raises capital, but it scales and creates inertia, which I needed to break through the market (ref. 3-Kaplan & Strömberg, *Financial Contracting Theory Meets the Real World*, NBER Working Paper, 2000).

Here is a breakdown of the stages:

1. Pre-seed and seed round:
    a. Purpose: financing MVP, market analysis and assembling a strong team
    b. Investor profile: family and friends, founders themselves, seed funds and angel investors
    c. Amount: relatively small sums, enough to fuel the startup until next round
    d. Valuation: low due to high risk and indefinite business model

2. Series A round:
    a. Purpose: optimize the product, market validation, customer acquisition and often revenue growth
    b. Investor profile: VCs and angels
    c. Amount: typically, around several million
    d. Valuation: increased, especially if traction has been proven

3. Series B and C rounds:
    a. Purpose: accelerate growth, expansion into new markets, possible internal changes and new hires
    b. Investor profile: VCs, equity funds and sometimes strategic corporate investors
    c. Amount: Around hundreds of millions
    d. Valuation: very high, risk reduced and scalability proven achievable
4. Later rounds and pre-IPO financing:
    a. Purpose: achieve market dominance and IPO pumping
    b. Investor profile: VCs, PEs, hedge funds and public investors
    c. Amount: large sums
    d. Valuation: highest possible

According to Kaplan & Strömberg (2000), only about 50% of seed-funded startups progress to Series A, and less than 10% eventually achieve a successful exit; this phased investment approach is designed to mitigate risk and allocate more capital only to companies showing strong early traction.

1.3) Deal flow process

This is the process adopted by VCs to search, discover, evaluate and select promising startups. This is a crucial and delicate process that allows VCs to navigate through thousands of new enterprises, which are born every day. Statistically, only 0.5 % of startups receive investment from larger entities such as VCs and funds, making the field competitive and hard to reach for (ref. 4- Kerr & Nanda, *Financing Young Firms*, Journal of Finance, NBER Working Paper, 2009). Here are the steps of the process:

- **Origination (sourcing**), this is where VCs and startups meet for the first time, it happens through professional networks of partners, startup events, demo days and so on. Sometimes VCs search actively for already established and valuable projects, and they are the ones reaching out to the entrepreneurs.

- **Screening**, After VCs gather different projects, they are filtered based on team background, market potential and alignment with fund's investment thesis.

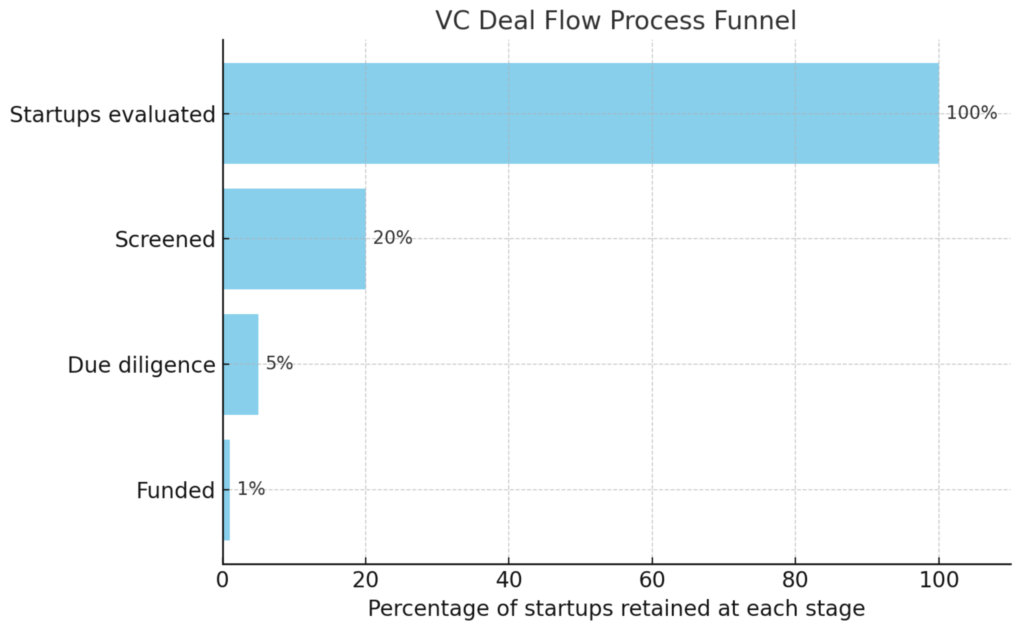
- **Due diligence**, once the new projects are selected, it is needed to analyze in depth (before investing) the target market, the business model, the team, the product and other financial metrics and forecasting.

- **Term sheet,** negotiation between the two parties, with clarification of terms and conditions such as equity percentages and invested sums along with voting rights and so on.

- **Closing**, the VCs team present the project to investors, then the capital is transferred to the startup.

- **Follow up**, the VC follows the projects by attending meetings, assisting with planning and monitoring KPI and financial metrics.

Along these main processes, many other dynamics take place for the VC to be successful. Expertise, professionalism, perseverance and diligence are qualities required when dealing with high-stake environments. These qualities are essential to carry out tasks such as the management of LP relationships and the investment committee, raising funds and creating value through strategic maneuvers. Good decision-making leads to good reputations, which is also an asset when exploiting professional networking and contacting funds entities.



1.4) How AI enhances VCs

In this sector, every valuable resource is taken into consideration and explored, and thus of course AI models. Venture Capital firms experience significant benefits from AI models which revolutionize their startup selection, evaluation and investment processes. The conventional practice of VC decision-making has depended heavily on experienced judgment together with network connections and traditional manual investigation methods. The essential role of human decision-making in the investment process remains still but AI provides data-based analysis to improve all stages of investment activities.

AI delivers its first major value through its capability to examine extremely large datasets at high speeds. The reality is that VC investors review numerous startups, but their restricted resources prevent them from conducting comprehensive evaluations. The automatic system of AI enables VC firms to discover promising startups by examining databases and websites and patent applications and social media and news sources. AI tools help venture capitalists detect investment opportunities which they would have otherwise missed.

AI can also create forecasts. Since AI models train on historical investment data, they can provide VCs with indications of the likelihood of a startup creating a commercially viable business or if it will likely fail. This information can assist VCs in determining which opportunities to prioritize further discussion with a founder. However, despite models not being 100% accurate, they can also provide a second opinion which can be a reference for what should be pursued in detail or may act as a new red flag that did not exist previously. In summary, while AI does not replace the insight and experience of VCs, it can aid the decision-making process for VCs by enabling them to make better, faster and more scalable investment decisions in an evolving, competitive marketplace (ref. 5-OECD, *Venture Capital Investments in Artificial Intelligence*, 2023).

Chapter 2

AI models

Artificial intelligence is a wide and interdisciplinary sector of computer science, which focuses on developing systems that can perform tasks usually done by human intelligence. So, tasks such as language and imagine recognition, reasoning, problem solving and so on. One of the main advantages of AI models is their capability of improving over time without human intervention.

From a historical perspective, AI is not recent; in fact, since the 1950s with Alan Turing, the concept of intelligent machines begun to be considered within professional environments. Later in the 1960s, Jhon McCarthy, during the Dartmouth conference defined the term "Artificial Intelligence" (AI). From then on, general interest in the subject declined until the 1990s when the concept of Machine Learning, a data-driven algorithm, was introduced. In 1997 IBM's Deep Blue was able to defeat chess world champion Garry Kasparov. In the late 2000s then, along with the internet and digital transformation, deep learning (DL) had an explosion of interest. It is inspired by the human brain's neural architecture and led to breakthroughs of CNNs and RNNs (which we will see later). In the 2020s AI reached a new level of capability with the adoption of foundation models, which are large and trained models that can be fine-tuned to a variety of tasks, worth mentioning is the widely known and used Chat GPT, launched by OpenAI.

What made AI the powerful tool as we know it today was the advance in data-processing technology, as AI models require analyzing large amount of data to perform. In particular, GPUs (Graphic Processing Unit) mainly provided by the notorious company Nvidia, which in the 2000s was backed by VCs such as Sequoia Capital and Sutter Hill Ventures.

In the last years AI technology has been used across different industries driving innovation and efficiency. In the healthcare sector, doctors use AI tools to diagnose diseases with more accuracy and speed. A private institute in Texas, AI is used as a teacher with incredible results (its students score in the 2% nationally). Large models such as Chat GPT have transformed content creation and customer service. In manufacturing, it is possible (using sensors and AI models) to enhance predictive maintenance and process automation. In finance, algorithms are now used to detect fraud and predict market trends. Thus, business processes are also becoming a key target of AI adoption, and we are likely to witness a widespread transition from traditional digital systems to AI-integrated operations in the near future.

2.1) Classification of the main AI models

Artificial Intelligence is often treated as a single concept, but it includes a wide range of models, each with its own logic, architecture and purpose. What they all have in common is the goal of replicating, in different ways, some forms of human intelligence, such as recognizing patterns, understanding language, making predictions or solving problems.

To make sense of this vast field, it helps to group AI models into three main categories:

- machine learning (ML)

- deep learning (DL)
- natural language processing (NLP)

These aren't rigid divisions, and in many real-world applications they overlap, but they are useful for understanding how different models work and what they are used for.

Machine learning is the foundation. It refers to algorithms that learn from data, without being explicitly programmed for every single task. Instead of writing a fixed set of rules, the model is trained on historical data and then uses that knowledge to make decisions or predictions on new, unseen data. The most classic examples include regression, classification, decision trees and support vector machines. These models are often used when the dataset is structured and well-organized, and when we need a certain level of interpretability.

Deep learning, on the other hand, takes inspiration from the human brain and uses structures called neural networks, made up of multiple layers. These models are much more complex and powerful than traditional ML, and they can identify subtle patterns in huge amounts of unstructured data, like images, audio or videos. They require more data and more computing power, but their performance is often superior. Within DL we find models like convolutional neural networks (CNNs), used for image recognition, or recurrent neural networks (RNNs), used for sequences like time series or text.

Natural language processing is a field on its own, focused on the interaction between machines and human language. NLP models are built to understand, interpret and generate human language, both written and spoken. Earlier NLP systems were rule-based or statistical, but modern NLP has become part of the deep learning revolution. Thanks to transformer architectures and large datasets, today's models are capable of tasks such as sentiment analysis, translation, summarization, and even holding a conversation.

Each of these categories developed and gained popularity in response to specific needs. Machine learning rose when we had more structured data than clear rules. Deep learning took over when data became too complex for traditional methods. NLP exploded once the internet made language data abundant and transformer-based models made it possible to process it effectively.

Understanding the differences between these families of models is important, because depending on the problem, whether it's classifying startups, analyzing documents, or predicting growth, some models are better suited than others.

2.2) NN, clustering and predictive models

Within the broad domain of artificial intelligence, a few model families stand out as foundational tools for solving complex tasks related to pattern recognition, decision-making, and data abstraction. This section focuses on three core categories: artificial neural networks (ANNs), clustering algorithms, and predictive models. Each is based on different mathematical principles, serves distinct purposes, and fits specific data types and problem structures (Ref. 6-Chollet, F. *Deep Learning with Python,2021*).

Neural Networks

Firstly, artificial neural networks (ANNs) are computational systems designed to recognize patterns and learn from data in a way that mimics, in abstract form, how biological neurons operate. They are the foundation of most modern deep learning approaches.

At the most basic level, a neural network is composed of layers of nodes (also called artificial neurons):

- An input layer that receives the features from the dataset (pixel values, numerical inputs, word embeddings).
- One or more hidden layers where intermediate representations are learned.
- An output layer that produces the final prediction (a class label, a numerical value).

Each connection between nodes has a weight, a trainable parameter that defines the strength of the signal passed between neurons. A neuron computes a weighted sum of its inputs, adds a bias and passes the result through a non-linear activation function. Common functions include:

- Sigmoid: compresses values between 0 and 1. Useful for binary classification.
- ReLU (Rectified Linear Unit): $f(x) = \max(0, x)$, widely used for its simplicity and performance.
- Tanh: scales outputs between –1 and 1, sometimes used in RNNs.

The network learns by comparing its predictions with the actual targets using a loss function. It then updates the weights through backpropagation, an algorithm that calculates the gradient of the loss with respect to each weight using the chain rule of calculus.

The actual updates are made using gradient descent (or its variants like Adam, RMSprop), which iteratively moves the weights in the direction that minimizes the loss.

A feedforward network passes information in one direction, from input to output, without loops. These are used for simple classification and regression tasks.

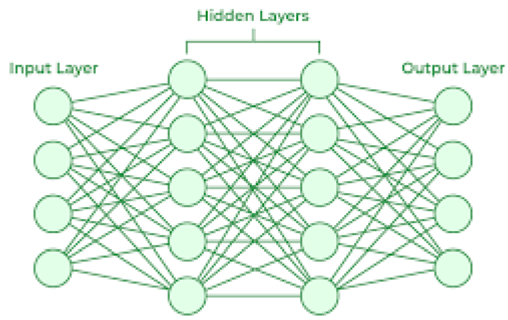
A recurrent neural network (RNN), on the other hand, allows connections that form cycles. This makes them ideal for sequence-based data as they can maintain a memory of previous inputs. Variants like LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) address the problem of vanishing gradients and enable long-term dependency learning.

Convolutional Neural Networks (CNNs) are optimized for spatial or grid-like data, such as images or audio spectrograms. Instead of using fully connected layers, CNNs apply convolutional filters that scan across the input to extract local features (like edges, textures). A CNN usually includes:

- Convolutional layers: apply kernels to input volumes.
- Pooling layers: down sample feature maps to reduce dimensionality.
- Fully connected layers: perform final classification based on extracted features.

CNNs reduce the number of parameters drastically compared to dense networks, making them efficient and highly effective in image-related tasks.

The key strength of neural networks lies in their ability to learn hierarchical representations. The lower layers of the network learn basic features (edges, textures), intermediate layers learn patterns (shapes, structure), and higher layers learn abstract concepts (object identity, sentiment, intent). This feature abstraction allows neural networks to outperform traditional models on complex, high-dimensional datasets.



Despite their power, neural networks are often seen as black boxes, meaning it can be difficult to interpret how or why a particular decision was made. Additionally, they require:

- Large and labeled datasets
- High computational power (typically GPUs)
- Careful regularization (dropout, early stopping) to avoid overfitting

Recent developments in explainable AI (XAI) aim to address some of these concerns, but transparency remains a challenge.

Clustering

While supervised learning methods rely on labeled datasets to guide their predictions, clustering belongs to the family of unsupervised learning. It's designed for a different purpose: to find structures in data where no explicit labels exist. In simple terms, clustering algorithms look for patterns, similarities, or proximity between data points, and group them into clusters, each ideally representing a natural segment within the dataset.

This type of approach is especially useful when we want to explore large datasets, identify hidden groupings, or simplify data for further analysis. The key assumption behind clustering is that similar items belong together, but the definition of “similar” depends on the method chosen and the context of the data.

Most clustering algorithms work by analyzing the distance or similarity between data points in a multi-dimensional space. The closer two points are (based on a distance metric like Euclidean or cosine), the more likely they are to belong to the same group.

Importantly, these models do not know in advance how many clusters exist, what shape they take, or what each cluster should mean. The model's job is to discover the underlying structure, and it's up to the analyst to evaluate whether that structure makes sense for the problem at hand.

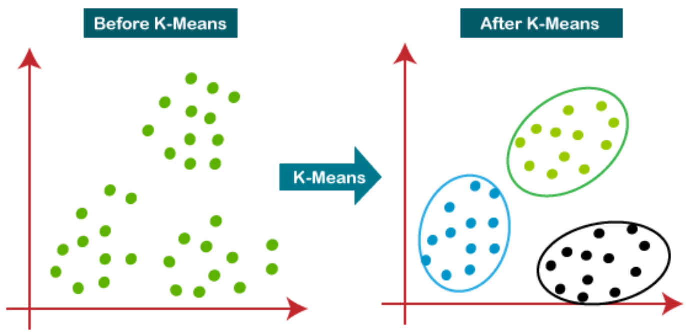
The most widely known clustering algorithm is K-means. It's relatively easy to implement and works well when the clusters are compact and well-separated.

Here's how it works:

1. We choose the number “K” (the number of clusters).
2. The algorithm randomly places K centroids in the space.
3. Each data point is assigned to the nearest centroid.
4. The centroids are recalculated as the average of the assigned points.
5. Steps 3 and 4 are repeated until the centroids no longer move significantly.

This process minimizes the intra-cluster variance, for example it tries to keep points within the same cluster as close to each other as possible. The result is a set of clusters, each defined by its centroid

However, K-means has a few well-known limitations: It assumes clusters are spherical and equally sized, it requires the user to know the number of clusters (K) in advance and it's sensitive to outliers and initial conditions.



In practice, these issues can sometimes be mitigated by repeating the algorithm with different initializations, or by using methods like the elbow method to estimate the optimal K.

For datasets where clusters are irregularly shaped, or where noise and outliers are common, density-based algorithms are often a better choice. One of the most popular is DBSCAN (Density-Based Spatial Clustering of Applications with Noise).

DBSCAN groups together points that are densely packed and separate them from areas of low point density. It does not require specifying the number of clusters in advance, and it can automatically label sparse regions as noise. This makes it more robust in messy or real-world data, where clusters aren’t clearly defined.

Key parameters include:

- ε (epsilon): the maximum distance between two points for them to be considered part of the same neighborhood.
- MinPts: the minimum number of points required to form a dense region.

DBSCAN works well with non-globular clusters and is ideal when we suspect there may be noise or anomalies in the data.

Another important approach is hierarchical clustering, which builds a tree-like structure (called a dendrogram) that shows how clusters are nested within each other.

There are two main strategies:

1. Agglomerative(bottom-up): Each data point starts as its own cluster, and pairs of clusters are merged step-by-step based on similarity.
2. Divisive(top-down): All points start in one large cluster, which is then recursively split into smaller groups.

This method does not require choosing the number of clusters upfront. Instead, you can “cut” the dendrogram at the desired level of granularity to extract the clustering result. It is particularly useful when we want to explore the data at multiple resolutions (ref. 7-Xu & Tian, *A Comprehensive Survey of Clustering Algorithms*, Annals of Data Science, 2015).

There is no universally "best" clustering algorithm. The choice depends heavily on:

- The shape and distribution of data.
- Whether the number of clusters is known or needs to be discovered.
- The tolerance of noise and outliers.
- The scalability needs (large datasets may challenge hierarchical methods).

Because clustering doesn’t rely on predefined labels, evaluation is often subjective and requires a good understanding of the context. That said, there are some internal metrics (like silhouette score or Davies-Bouldin index) that help compare cluster quality quantitatively.

Clustering shines in situations where patterns are not obvious, or when the goal is discovery rather than prediction. It can reduce dimensionality, expose subgroups, and offer a new lens for interpreting the data.

However, because it's unsupervised, clustering models can be difficult to validate, and their output may depend heavily on input assumptions and parameters. They’re not designed to predict future outcomes, but to reveal structure, which makes them a valuable first step in exploration of data analysis.

2.3) LLM and generative models

(ref. 8- Bommasani et al., *On the Opportunities and Risks of Foundation Models*, Stanford CRFM, 2021).

Over the last few years, a new class of models has dramatically expanded the capabilities of AI systems, Large Language Models (LLMs) and more broadly, generative models. These models don’t just classify, cluster, or predict. Instead, they can generate new content, such as text, images, code, or even music, by learning from massive datasets and identifying patterns that are not just statistical, but also creative in nature.

Large Language Models are neural networks trained on enormous corpora of text. They belong to the transformer architecture family, introduced in 2017 with the paper “Attention Is All You Need.” Unlike previous models that processed text sequentially (like RNNs), transformers use a mechanism called self-attention, which allows them to consider all parts of a sentence at once. This makes them extremely efficient at modeling long-range dependencies in language.

Modern LLMs, such as GPT (Generative Pre-trained Transformer), BERT (Bidirectional Encoder Representations from Transformers), or T5 (Text-to-Text Transfer Transformer), are trained in two steps:

1. Pre-training: The model learns general language patterns from a vast dataset (internet, books, Wikipedia, etc.).
2. Fine-tuning: The model is adapted to specific tasks (like summarization, translation, question answering) using curated datasets and human feedback.

What makes LLMs powerful is their generalization ability. A well-trained model can perform a wide variety of tasks with little to no further training, what's known as zero-shot or few-shot learning.

LLMs are a subset of a broader field called generative AI, which focuses on producing novel content. Other examples include:

- GANs (Generative Adversarial Networks): Two networks compete, one generates data, the other evaluates it. Used to generate realistic images or videos.
- VAEs (Variational Autoencoders): Models that learn efficient data encoding and can sample from a latent space to create new outputs.
- Diffusion models: Recently popular for high-quality image generation. These models learn to denoise data in steps, effectively learning the distribution of complex images.

Generative AI can mimic human creativity, but it does so by learning statistical patterns at an immense scale. It doesn't "understand" language or images in a human sense; it models probability distributions and generates outputs that are statistically likely to follow a given input.

Generative models are extremely versatile. In the language domain, they can generate human-like text, translate languages, summarize long documents, answer open-ended questions and write code or emails.

In other domains, they can: create synthetic images or audio, generate 3D models or music, design new molecules in biochemistry, simulate data for training other models.

These models are often used via APIs and deployed as part of larger pipelines. In many scenarios, they serve as a starting point that is then refined or reviewed by a human user.

Despite their impressive output, generative models come with caveats:

- They are data-hungry and computationally expensive to train.
- Their results can be unpredictable or hallucinatory (confidently incorrect).
- They often lack transparency and cannot explain why a particular output was chosen.
- They carry risks related to bias, misinformation, and ethical misuse.

As we will discuss later in Chapter 6, these limitations raise important questions when deploying such models in high-stakes or regulated environments.

2.4) AI limits

While artificial intelligence has achieved remarkable results in recent years, it is important to understand that AI is not magic, nor is it universally applicable. Every model operates within certain boundaries, and knowing these limits is crucial for responsible use and realistic expectations (ref. 9-Burrell, *How the machine 'thinks': Understanding opacity in machine learning algorithms*, Big Data & Society, 2016).

AI models do not "understand" the world, they learn patterns from training data. If that data is incomplete, biased, or not representative, the model will reflect those same issues in its output. For example, if a sentiment analysis model is trained mostly on English social media from North America, it may struggle to interpret slang or cultural references from other regions. In healthcare, if a model is trained mostly on data from men, its diagnostic accuracy may be significantly lower for women.

Many models perform extremely well on training data but fail when exposed to new, unseen inputs. This is known as overfitting. It occurs when the model learns not just the general patterns but also the noise or specific quirks of the training set.

To combat this, practitioners use cross-validation, regularization, and dropout layers (in neural networks), among other techniques. Still, generalization remains a central challenge in real-world applications.

Modern models, especially deep learning and transformers, are often referred to as black boxes. They contain millions or even billions of parameters, and it's nearly impossible to trace how a specific decision was made.

This raises problems in sectors that require accountability, such as:

- Finance (loan approval)
- Law (judicial support tools)
- Medicine (clinical diagnosis)

Efforts to address this include the growing field of Explainable AI (XAI), which seeks to develop models and interfaces that provide understandable justifications for outputs. However, this is still a developing area and not yet standard in most applications

AI systems can amplify societal biases if those biases are present in the training data. This has been observed in areas like facial recognition (lower accuracy for non-white individuals), job recommendation systems, predictive policing tools.

Without deliberate bias mitigation strategies, AI can unintentionally reinforce discrimination or inequality.

Other critical concerns are model misuse (deepfakes, disinformation), data privacy and surveillance, job displacement due to automation and over-reliance on automated decisions without human oversight.

These are not just technical limitations, they intersect with legal, economic, and social considerations that must be addressed holistically.

Chapter 3

Application of AI in VCs

Now that we've explored the main families of AI models and how they work, it's time to link that knowledge to the real world. More in depth, how they can enhance the operations of VC firms.

Venture Capitals have always operated in uncertain and competitive environments, where the critical difference is given by speed, access to information, and quality of judgment. Usually, investment decisions relied heavily on human intuition, personal networks, and manual analysis. But, as the volume of available data has exploded and markets have become more saturated, this approach has started to show its limits, or better, has started to show room for further improvements.

Artificial intelligence offers new tools to augment, accelerate, and refine the decision-making processes that stand at the core of venture investing. From scouting early-stage startups to assessing founding teams and forecasting future growth, AI can support VCs at every step, always if applied correctly.

In this chapter, we'll look at the concrete ways AI is already being used in this market. Each section will focus on a specific area of application (scouting, valuation, prediction, natural language processing) and explore how AI models are changing the rules of the game. These are not just theoretical possibilities: many leading firms are already adopting these tools and integrating them into their operations (ref. 10 -Scabbio, *Exploring the Adoption of Artificial Intelligence in Venture Capital*, 2024).

This is where the abstract models introduced in Chapter 2 start to take on strategic value, and where we begin to see how data science and investment logic can be combined into a new way of thinking about venture capital.

3.1) Scouting automation

This is the first and earliest stage of venture capital activity. The challenge now is no longer finding a promising startup, but it is navigating through the overwhelming abundance of them. Every year, tens of thousands of new ventures are created across sectors and locations. For VC firms, the bottleneck is not a lack of opportunities, but the ability to detect high-potential startups early (before competitors do) and before the signal is diluted by noise (ref. 11-Ozince & Ihlamur, *Automating Venture Capital: Founder assessment using LLM-powered segmentation, feature engineering and automated labeling techniques*, arXiv, 2024).

Traditionally, sourcing was based on founders reaching out, personal referrals, networking events, and curated pitch meetings, but in the digital era, much of startup's activity leaves digital traces across a wide ecosystem of platforms: LinkedIn, GitHub, Product Hunt, Crunchbase, Twitter (now X), Medium, subreddits, angel platforms, and accelerator databases. For a human analyst, tracking even a small fraction of this digital flow is practically impossible, or it requires huge amounts of time and effort.

This is where AI-powered scouting enters the picture, not to replace human judgment, but to support it. AI systems can automate the monitoring, classification, and prioritization of early-stage startups by consuming real-time and unstructured data from multiple sources, than transforming it into structured insights. Natural Language Processing (NLP) is used to extract entities, keywords, and contextual cues from unstructured sources such as founder interviews, tech blogs, or news sites. Models such as spaCy, BERT, or DistilBERT can easily identify and classify references to companies, products, technologies, locations, or funding events.

Web scraping pipelines, combined with models like Named Entity Recognition (NER), can allow automated systems to build structured startup profiles from raw HTML, tweets, job boards, or GitHub repositories. This enables detection of freshly incorporated startups, hiring growth signals, or MVP launches before they land on the mainstream media.

Clustering and unsupervised learning algorithms can detect thematic patterns across large groups of firms. For example, if multiple startups begin to emerge using similar keywords, clustering algorithms can signal the formation of an emerging sub-market.

Graph-based models, often built on libraries like NetworkX or Neo4j, are used to analyze relationships between founders, investors, and companies. These can flag recurring founder-investor connections, “power clusters” of talent (for example ex-Palantir or Y Combinator employees), or co-founding networks that tend to start successful ventures.

Predictive scoring models, based on classical ML or deep learning (MLPs or transformer encoders) can learn to assign likelihood scores to startups based on historical data.

Inputs might include team composition, product type, geographic region, traction metrics and web activity. The model’s output is often a probability of success, or a suggestion for deeper review. These models can be retrained periodically to reflect shifting investment theses, macro trends, or firm-specific strategies.

A modern AI-assisted scouting pipeline might include operations like:

- Data ingestion: Collection of live data streams from APIs (Crunchbase, AngelList, LinkedIn) and scrapers for news, blogs, and registries.
- Data enrichment and NLP tagging: Using models like Flair, AllenNLP, or OpenAI embeddings to extract meaningful metadata (industry, founder background, use case, keywords).
- Similarity search: Applying vector-based models (FAISS) to compare new startups with previous successful cases or target profiles.
- Ranking and filtering: ML algorithms trained on prior investment outcomes to sort and prioritize leads.
- Analyst interface: Surfaced results are ranked by relevance and uncertainty, with transparency indicators showing the reason why a startup was selected.

These systems don’t generate final decisions, they analyze overlooked signals, helping analysts focus their time and effort on the most promising leads.

The main advantage of AI-driven scouting is the ability to scale the top of the funnel without losing relevance. Specifically:

- Increased volume: Systems can monitor tens of thousands of entities in parallel.
- Faster reaction times: Early signs (a prototype launch or a strategic hire) can be flagged within hours instead of weeks.
- Unbiased coverage: AI can detect startups outside mainstream networks or elite circles, including in geographies or industries not typically monitored by VCs.
- Trend anticipation: Clustered keyword shifts or hiring waves can signal the rise of new categories, even before market analysts name them.

A clear example of how AI has revolutionized the scouting phase in venture capital is the drastic reduction in the time needed to screen and shortlist startups. Traditionally, investment analysts spent several days, sometimes even a full week, manually reviewing business plans, going deeper in founder networks, and cross-checking market data to create an initial shortlist. Today, AI-powered tools can automate large parts of this process, leading to remarkable time savings and higher precision.

For instance, a recent study by CBS Copenhagen Business School (Jhon, Moser - CBS Copenhagen Business School, "AI in Private Equity and Venture Capital" 2022) documented that advanced data mining algorithms now enable venture capital firms to complete an initial evaluation of a startup in just one or two days, compared to the typical five to seven days required with traditional methods. This acceleration not only improves operational efficiency but also provides a strategic advantage: investors can approach promising founders faster and with more accuracy, gaining a head advantage in highly competitive markets.

Furthermore, reports show that about 42% of global VC firms already use AI tools for deal sourcing and scouting, and this figure rises to 75% among top-tier funds, underlying its capabilities. These tools scan diverse sources, including startup databases, academic publications, patent filings, and social media sentiment, offering a multilateral view of emerging companies that would be impossible to gather manually in such a short timeframe.

Despite these advantages, AI models for scouting face several limitations, the main being:

- Data quality: Many promising early-stage startups operate in stealth mode or do not maintain a strong online presence, making them invisible to digital-first methods.
- Bias reinforcement: Models trained on historical VC data may inherit existing biases (favoring male founders, Silicon Valley hubs, or certain educational backgrounds), unless counterbalanced by explicit fairness techniques.
- False positives: Hype-driven signals (media buzz, hiring bursts) can overinflate certain startups' perceived potential.
- Lack of qualitative context: AI can score a team's past experience but not necessarily grasp its internal dynamics, vision alignment, or long-term resilience.

In this sense, AI is best seen as a first filter, not a gatekeeper. It reduces the noise but leaves the final judgment to human analysts.

AI-powered scouting is transforming how VC firms discover new opportunities and startups. It's not about fully automating investment decisions, but about scanning broader and more accurate, reacting faster, and acting earlier, gaining competitive advantage, especially in an environment where the timing and visibility of information can create a critical edge.

3.2) valuation of market, team and traction

In venture capital, the valuation of a startup is mostly uncertain, especially in early stages, where historical data is limited and the potential is still speculative. To diversify risk and maximize expected returns, VCs rely on a detailed evaluation of three key dimensions: market opportunity, founding team quality, and traction indicators. Usually, these evaluations are subjective, and intuition driven. Nonetheless, the integration of Artificial Intelligence now enables a systematic, data-driven approach to assess these parameters with greater precision, speed, and scalability (ref. 12-Montanaro et al., *Venture capital investments in artificial intelligence*, Journal of Evolutionary Economics, 2024).

Market Valuation: AI in Opportunity Sizing and Competitive Mapping

A fundamental rule of venture investment is that even the best team and product is going to fail in a small or stagnant market. Thus, understanding the Total Addressable Market (TAM), Serviceable Available Market (SAM), and Serviceable Obtainable Market (SOM) is critical (ref. 13-Fang, Tao & Li, *Anchoring AI Capabilities in Market Valuations: The Capability Realization Rate Model and Valuation Misalignment Risk*, arXiv, 2025).

AI models assist in market valuation by:

- Web scraping and NLP: AI can process thousands of market reports, blogs, regulatory filings, and news articles to estimate market size, segment evolution, and customer demand patterns.
- Clustering techniques: Market entities can be grouped by industry codes, business models, or customer personas to reveal under-served segments or emerging niches.
- Trend forecasting models: Time-series analysis and sentiment mining on platforms like Google Trends, Twitter, or industry forums allow prediction of trend inflection points and early-market signals.
- Competitive intelligence tools: Startups can be benchmarked using AI-powered tools that analyze their public footprint (web traffic, app usage, SEO rank) to model their competitive positioning.

For instance, platforms such as Crunchbase and Pitchbook have been integrated with ML pipelines that extract competitive signals (funding raised, acquisition activity, hiring trends) to dynamically adjust market opportunity scoring.

Team Evaluation: From Résumé Parsing to Dynamic Scoring

The team's ability to execute the business model under uncertainty is often considered the most qualitative factor, but AI is making it more quantifiable.

AI can assist in:

- Automated résumé analysis: NLP models can parse educational and professional backgrounds, assessing alignment with industry verticals and past success patterns.
- Social graph analysis: Using LinkedIn or public data, graph neural networks (GNNs) can assess the founders' network centrality, reachability to strategic partners, or access to follow-on investors.
- Track record prediction: By comparing with a database of past founders (those who exited successfully or failed), classification models can output a risk score or probability of success.
- Video/audio sentiment analysis: AI tools can analyze pitch recordings, detect confidence, consistency, leadership cues, and linguistic markers associated with high-performing entrepreneurs.

Moreover, some VC firms use psychometric AI tools to measure qualities like resilience, adaptability, and problem-solving mindset, extracted from behavioral data or linguistic analysis of communications.

Traction Analysis: From Descriptive KPIs to Predictive Growth Models

Startup traction is perhaps the most quantitative component of the three. It includes metrics such as Monthly Recurring Revenue (MRR), user base growth rate, churn rate, Customer Acquisition Cost (CAC), Lifetime Value (LTV), Daily Active Users (DAU), and more.

AI helps evaluate traction in multiple dimensions:

- Real-time data ingestion: APIs and AI tools integrate with accounting systems, CRMs, or analytics platforms (Stripe, Mixpanel) to collect updated performance data.
- Time-series modeling and forecasting: LSTM (Long Short-Term Memory) neural networks can model seasonality, detect anomalies, and forecast growth patterns.
- Unit economics benchmarking: AI compares the startup’s metrics with industry benchmarks (CAC/LTV ratios), flagging outliers or unsustainable business models.
- Cohort analysis and clustering: Users can be segmented by usage patterns or value over time, identifying whether traction is driven by high-retention users or one-time spikes.

Importantly, AI helps detect the so-called "vanity metrics" (inflated downloads or social media followers) by horizontally referencing with engagement data, churn, or conversion rates, enhancing the reliability of traction signals.

Rather than treating market, team, and traction as isolated assets, AI models, especially ensemble models, can synthesize all variables into a unified and standardized scoring system. This allows for:

- Startup ranking based on a weighted function of market growth rate, founder credibility score, and normalized traction KPIs.
- Investment readiness index: calculating a probabilistic metric of success or readiness for next funding round, based on similar historical cases.
- Outlier detection: flagging startups that may have weak current KPIs but strong similarity with unicorns at earlier stages.

Tribe Capital, for example, uses a proprietary model called the “Data-Driven Product-Market Fit Signal”, which quantifies user love and network effects using raw data from product analytics (ref. 14-Potanin et al., *Startup success prediction and VC portfolio simulation using CrunchBase data*, arXiv, 2023).

Nonetheless there are some limitations, while AI enhances decision-making, it must be acknowledged that: historical bias in training data (predominantly male founding teams) may lead to exclusion of underrepresented founders, data quality issues may distort predictions, especially for early-stage startups with limited operational history and the interpretability of models is critical, especially when justifying investment decisions to LPs or board members.

3.3) Successful predictive algorithms: unicorn clustering, fundraising pattern

The application and use of predictive algorithms in Venture Capital is gaining popularity as the firm’s goal to identify high-potential startups earlier and with higher accuracy. Traditional evaluation methods, while full of human insight, are limited in scalability and often suffer from cognitive bias. AI-based predictive models provide an additional, data-driven view to detect early signals of success, particularly in recognizing "unicorn-like" patterns and forecasting fundraising outcomes. In this section, we explore how AI contributes to building predictive frameworks that can identify potential unicorns and anticipate funding processes.

Unicorn Clustering: Pattern Recognition Across Historical Data

A unicorn is defined as a privately held startup valued at over $1 billion. These startups often follow similar growth paths or exhibit shared characteristics in their early stages. AI models, especially clustering algorithms and ensemble classification systems are able to uncover these commonalities. How it works, by step:

1. Data aggregation: Historical data on successful unicorns is collected, including metrics like revenue growth rate, DAU/MAU ratios, founding team composition, initial market size, and funding history.
2. Feature selection: Using dimensionality reduction techniques (PCA, t-SNE), relevant predictors are extracted from high-dimensional datasets.
3. Clustering algorithms: Models like K-means, DBSCAN, or Gaussian Mixture Models are trained to group startups into clusters based on similarities to known unicorns.
4. Classification layer: Once clusters are formed, a supervised learning model can be trained to predict unicorn probability for new startups entering the pipeline.

These models allow VCs to uncover startups that look alike that may not yet have traction but show similar profiles to past unicorns in terms of product growth, market velocity, and founder behavior (ref. 15-Cao et al., *Using Deep Learning to Find the Next Unicorn: A Practical Synthesis*, IJCAI FinNLP/Muffin Workshop, 2023).

Predictive Modeling of Fundraising Success

Another application of AI is in forecasting the probability and timing of future funding rounds, which is critical for portfolio planning and deal flow prioritization. Fundraising patterns often follow predictable movements, based on metrics like burn rate, monthly growth, and external investor interest. AI-driven fundraising prediction involves:

- Survival analysis models to estimate time-to-next-funding
- Gradient boosting machines (GBMs) to predict the likelihood of receiving Series A or B within a certain time frame
- Sequence modeling using RNNs or Transformer-based models on event data (e.g., press mentions, hiring surges, product launches)
- Sentiment analysis on public mentions and founder communication, which correlates with investor interest

These predictions can inform internal prioritization, for example, if a model forecasts an 85% likelihood of Series A fundraising in the next 6 months, the VC may move pre-emptively to secure a lead position or offer bridge financing (ref. 16-Calafiore et al., *Survival and Neural Models for Private Equity Exit Prediction*, arXiv, 2019).

Many VC firms are employing ensembles of predictive models, blending different algorithmic approaches to improve robustness. These include tree-based models for tabular KPI data, deep learning models for sequential or behavioral data and Bayesian models for uncertainty estimation and updating probabilities as new data is acquired.

Some firms, such as SignalFire and Tribe Capital, have built AI engines that integrate structured data (financials, KPIs) with unstructured signals (news, web traffic, founder behavior) to produce dynamic scoring dashboards.

While predictive algorithms offer powerful insights, they are not infallible:

- Data quality and coverage can skew predictions, especially for early-stage or stealth-mode startups
- Overfitting risks arise if models are too tightly tuned to historical unicorns, missing novel patterns
- Bias reinforcement is possible if models trained on skewed historical data favor certain demographics or geographies

3.4) NLP use for speech and documents

In the context of venture capital, NLP tools have become essential for handling the massive volume of unstructured textual data that flows through emails, pitch decks, blogs, news articles, financial reports, and even spoken conversations.

NLP transforms this unstructured data into structured insights by performing tasks such as named entity recognition (NER), topic modeling, sentiment analysis, summarization, and semantic search. These capabilities empower venture capital firms to extract meaningful information automatically, reducing the time and resources needed for manual review and increasing scalability in operations (ref. 17-Ozince & Ihlamur, *Automating Venture Capital: Founder assessment using LLM-powered segmentation, feature engineering and automated labeling techniques*, arXiv, 2024).

Many VCs receive pitch information through verbal channels, such as video calls, voice messages, and recorded meetings. NLP models, when paired with Automatic Speech Recognition (ASR), can transcribe spoken words into text and apply further NLP techniques to analyze the content. This helps in:

- Extracting key points and action items from meetings.
- Evaluating tone and confidence of founders during pitch presentations.
- Automatically storing and indexing conversations for future reference and compliance purposes.

Modern tools such as Otter.ai or Microsoft Azure Speech Services offer transcription coupled with sentiment and keyword analysis, allowing VCs to assess not just what is said, but how it is said.

One of the most time-consuming activities for VC analysts is the review of documentation provided by startups. These include pitch decks, business plans, market reports, cap tables, and legal contracts. NLP systems can:

- Summarize long documents, highlighting key insights (competitive advantage, traction metrics).
- Extract specific elements, such as financial projections or team background
- Cross-reference claims made in the pitch with external sources like Crunchbase, LinkedIn, and news media.

For example, semantic search engines powered by NLP can scan through thousands of PDF documents to detect recurring red flags or missing elements, accelerating early screening. Named entity recognition models can also identify key players (individuals, companies, locations) and link them to existing knowledge bases to evaluate credibility and network strength.

Startup-founders and VCs exchange hundreds of emails throughout the scouting and due diligence process. NLP models can: analyze communication patterns to detect responsiveness and engagement, identify sentiment trends over time, group and tag messages according to themes (fundraising, product milestones, competition, etc.). This can be valuable to prioritize which relationships require follow-ups, which conversations indicate strong signals of potential, and which threads show signs of concern (ref. 18-Tran et al., *Parsing Speech: A Neural Approach to Integrating Lexical and Acoustic-Prosodic Information*, arXiv, 2017).

The main advantage of NLP in VC processes is automation at scale. VCs can process and interpret far more data than previously possible, enhancing both deal flow efficiency and decision quality. It also mitigates human biases, offering a more neutral lens for analyzing language-based content.

3.5) Use cases of these tools

The theoretical and technical applications of Artificial Intelligence within venture capital processes find concrete manifestation in real-world use cases. Several firms and platforms have already integrated AI into their core decision-making and operational workflows, each with unique objectives and outcomes. This section examines how Zebra, Tribe Capital, and Craft AI utilize AI to improve scouting, evaluation, and portfolio management, serving as examples of how AI adoption is reshaping the VC landscape.

1. Zebra: Data-Driven Matching Between Startups and VCs

Zebra is a platform that aims to democratize access to capital for early-stage startups by connecting them to the right investors through AI-driven matching. It acts as an intermediary between thousands of startups and venture funds, using data science to streamline the scouting and selection process. Key AI functionalities:

- Startup profiling and clustering: Zebra uses NLP to extract relevant features from pitch decks, websites, and applications. Startups are then clustered based on product category, business model, and traction level.
- Investor matching algorithm: a recommender system, trained on successful startup-investor matches, scores compatibility between a startup and various VCs based on investment thesis, stage preference, geography, and past deals.
- Continuous learning: the platform refines its recommendations using supervised feedback from investor engagement (click-throughs, meetings booked, actual investments).

Zebra demonstrates how AI can level the playing field for startups that may not have traditional network access, allowing capital to flow more meritocratically.

2. **Tribe Capital: The Quantitative VC**

Tribe Capital is a Silicon Valley VC firm that operates under a philosophy of "quantitative investing." It has developed a proprietary AI-powered framework to assess early-stage startups using product usage data and other structured signals. Key AI applications:

- Data-Driven Product-Market Fit (PMF) Signal: By integrating directly into a startup's internal data (usage metrics, retention, DAU/MAU), Tribe runs time-series models to detect organic engagement and virality.
- Founder and team evaluation: Tribe scores founders using past success metrics, social graph positioning, and pitch behavior analysis.
- Investment decision engine: Each startup is assigned a numerical score, combining growth KPIs, engagement depth, and market benchmarks. This score directly informs deal prioritization and due diligence depth.

Tribe Capital's method reduces subjectivity and standardizes early-stage assessment, increasing scalability and comparability across startups in the same cohort.

3. **Craft AI: Enabling Predictive Insights Across Portfolios**

While not a VC firm per se, Craft AI develops explainable AI infrastructure used by several investment firms to power insights across their portfolios. It focuses on interpretable machine learning, making it a key tool in regulated or risk-sensitive investment settings. Key AI features:

- Time-series forecasting: Craft AI enables prediction of revenue, user growth, or churn using interpretable models (decision trees, monotonic models).
- Explainability layer: Unlike black-box neural networks, Craft emphasizes "white-box AI" where decisions can be traced to individual features (revenue seasonality, pricing strategies).
- Anomaly detection: Portfolio companies are monitored in real time, and the system flags deviations from expected patterns in performance metrics.

Craft AI represents a crucial evolution toward transparent and compliant AI usage in financial decision-making, especially when models are shared with LPs or auditors.

Chapter 4

AI optimization of VCs processes

After exploring how Artificial Intelligence helps in individual decision points, such as startup scouting, founder evaluation, and fundraising prediction, this chapter shifts focus toward a broader, operational integration of AI inside of venture capital firms. Rather than looking at single and isolated applications, we examine how AI can optimize entire workflows, changing the internal mechanics of VC operations.

The goal is no longer just to improve accuracy in startup selection, but to enhance the efficiency, consistency, and scalability of the venture capital process. Starting from inbound deal triage to CRM automation, to due diligence to post-investment monitoring, VC firms are increasingly embedding AI into their daily infrastructure operations, often through bespoke tools or third-party platforms.

4.1) Automated Deal Flow Prioritization

One of the most promising AI uses in venture capital is in the automation and prioritization of the deal flow process. Normally, deal flow management is labor-intensive, requiring the VC team to manually screen and analyze thousands of startup proposals, filter based on fund thesis, and decide which deserves deeper evaluation. AI models can significantly enhance this workflow by automating the triage process, enabling funds to scale their outreach while maintaining analytical depth.

Machine learning models can be trained on historical investment data, such as pitch decks, market size, traction metrics, founding team background, and prior funding rounds, to classify incoming new startups based on their alignment with the fund's past successful investments. For example, classification algorithms can label startups as "high potential", "needs review" or "low priority" based on features extracted from all kind of sources that can be structured and unstructured (databases, web crawling, internal CRM notes, LinkedIn profiles, and even founder sentiment from pitch meetings).

This automation allows VCs to:

- **Reduce human bias** in the initial screening phase and avoid missing promising deals due to overload or unconscious preferences.
- **Focus analyst time on high-priority cases**, thus increasing operational efficiency and decision-making throughput.
- **Create a repeatable, data-driven methodology** that aligns with the fund's strategic investment thesis.
- **Track rejected startups** and flag them automatically in the future if updated data suggests improved viability or traction.

In implementation, Natural Language Processing (NLP) plays an important role. Tools like OpenAI embeddings or BERT-based models can extract semantic meaning from textual sources (pitch decks, founder emails, blog posts), allowing a deeper contextual ranking of each opportunity. Even more, tools like Affinity, Zint, or Craft Ventures' internal systems already use these approaches to improve pipeline quality.

Despite its benefits and advantages, this approach does not come without challenges. Models must be continuously updated to reflect evolving investment theses, new market trends, and changes in strategic aim. Moreover, the cold-start problem arises for startups with minimal digital footprints or operating in emerging niches, where models trained on past data might underperform (ref. 19-Arroyo et al., *Assessment of Machine Learning Performance for Decision Support in Venture Capital Investments*, IEEE Access, 2019).

4.2) Intelligent CRM and Relationship Management

Relationship management is a foundational pillar of venture capital prosperity. From maintaining contact with LPs, to managing relations with promising founders, co-investors, and advisors, venture capital firms rely heavily on their networks. Traditionally, this relationship-building process is supported by standard CRM (Customer Relationship Management) systems and personal memory. However, as deal volume gets higher and professional networks expand, this approach becomes less efficient and prone to information losses. AI-driven CRM systems offer a transformative solution.

Intelligent CRM platforms leverage machine learning and natural language processing to process and organize vast quantities of communication data (emails, meeting notes, call transcripts, and social media interactions). These tools can automatically:

- **Map relationship graphs** among founders, investors, and companies, highlighting key connectors and previously unknown links.
- **Analyze interaction history** to suggest when to follow up, who's slipping out of contact, or when a strategic introduction could be valuable.
- **Prioritize outreach** based on real-time activity or pattern recognition (similar profile to a previous success case).
- **Generate contextual insights** before meetings, such as a summary of past conversations, mutual connections, and relevant news.

Some AI-powered CRMs also integrate with platforms like LinkedIn, Crunchbase, and Twitter, extracting real-time updates about key people and startups. This enables VC professionals to always be well informed on job changes, fundraises, and market movements, without manual monitoring. Moreover, predictive analytics embedded in intelligent CRMs can signal which relationships are most likely to lead to investment opportunities, partnership potential, or portfolio synergies. For instance, systems like Affinity, Folk, and Salesforce Einstein are increasingly used by funds to not only track relationships but also to strategically manage them as an evolving pillar.

This approach transforms CRM from a passive database use to an active recommendation engine, making social capital measurable. It also reduces cognitive load on partners, enabling them to move on to the right relationships at the right time, supported by real-time data. Nevertheless, the success of these systems depends on data quality and user discipline. If meeting notes, emails or contacts are not systematically logged, the models will lack the training input necessary for effectiveness. Therefore, successful implementation often requires cultural adaptation alongside technical deployment (ref. 20- Lyu et al., *Graph Neural Network Based VC Investment Success Prediction*, arXiv, 2021).

4.3) Document Due Diligence Automation

Due diligence is a pivotal step in the VC investment process, taking in the in-depth assessment of a startup's financials, legal status, team, market size, competitive landscape, and technology. This phase, while crucial, is also time-consuming and repetitive, especially as VCs evaluate, as said before, hundreds of startups annually. Artificial Intelligence can significantly accelerate and improve this stage through automation tools specifically made to document analysis.

AI-powered due diligence models use Natural Language, Optical Character Recognition (OCR), and machine learning models to scan, extract, and interpret key information from a broad range of unstructured documents including:

- Business plans and pitch decks
- Cap tables and shareholder agreements
- Financial statements and forecasts
- Patents and technical documentation
- Legal contracts and compliance certificates

These systems can automatically find inconsistencies, missing documents, or unusual clauses, flagging them for human review. For example, an NLP engine can detect whether a startup's pitch claims conflict with figures in the financials, or whether the team's LinkedIn bios mismatch the roles listed in internal reports.

The benefits for VC firms are considerable:

- Speed: Document review time is drastically reduced, allowing faster deal execution.
- Scalability: Analysts can handle more deals simultaneously, without sacrificing depth.

- Risk reduction: AI can spot red flags that human reviewers may overlook due to time constraints or cognitive bias.
- Consistency: Standardized analysis reduces subjectivity across different partners or teams.

Some tools also provide automated generalized generation, allowing decision-makers to quickly get the essence of a deal without digging through extensive materials. Others offer document comparison insights, highlighting contractual differences across funding rounds or between comparable startups (ref. 21-Bai & Zhao, *Analyzing the Effectiveness of Artificial Intelligence in Due Diligence in Mergers & Acquisitions*, 2025).

4.4) Real-Time Portfolio Monitoring: KPIs, Analytics, and Forecasting

Once the investment is made, the role of the venture capitalist does not end. On the other hand, it transitions into a crucial phase: monitoring the performance and evolution of the portfolio companies. Startups are dynamic and high variance firms, maintaining real-time visibility into their development is essential for providing strategic positioning, anticipating potential issues, and planning possible exits. Artificial intelligence plays a transformative role in enabling this kind of continuous and intelligent oversight.

AI-driven analytics and predictive modeling tools allow venture capital firms to:

- **Track operational and financial KPIs in real time**, automatically aggregating data from diverse sources such as cash flow sheets, internal CRMs, project management tools and web or financial analytics platforms
- **Forecast future performance indicators** such as burn rate, revenue growth, user churn, or runway, through time-series forecasting models (ARIMA, LSTM) or Bayesian regressors
- **Detect weak signals or anomalies** that may indicate underlying issues or upcoming risks (e.g., a sudden drop in user engagement, delays in fundraising, missed milestones)
- **Generate dynamic dashboards and real-time visualizations** accessible to General Partners and Limited Partners, enabling standardized and transparent reporting across the portfolio

These capabilities make it possible to shift from a reactive to a proactive and predictive style of portfolio management. For example, AI systems can trigger alert notifications when a startup deviates from its key metrics, prompting immediate intervention or strategic reassessment by the team.

Moreover, AI enables comparative performance benchmarking across portfolio companies and versus external market standards, considering sector, stage, region, and size. This comparative view advantages decisions related to follow-on funding, operational support, or potential early exits.

Among advanced platforms supporting this functionality there is Visible.vc, Carta, Pulley, and Clairvoyant, while some firms have developed proprietary monitoring systems integrated with their own internal data pipelines.

Naturally, the effectiveness of such systems relies heavily on the quality, accuracy, and frequency of data shared by the startups. This is why many VCs now require standardized API-based data sharing or the mandatory use of specific software platforms by portfolio companies.

In conclusion, real-time portfolio monitoring powered by AI enhances the ability of venture capital firms to manage risk, provide timely support, and maximize portfolio value. It turns oversight from a manual, fragmented process into a continuous and intelligent strategic function underperforms (ref. 19 -Arroyo et al., *Assessment of Machine Learning Performance for Decision Support in Venture Capital Investments*, IEEE Access, 2019).

4.5) Risk Management and Anomaly Detection

As said in chapter 1, risk is inherent in venture capital industry, where high failure rates are balanced by the potential for outsized returns. Traditional risk management in VC has relied heavily on qualitative and human judgment, board oversight, and periodic reporting. However, in an increasingly complex and fast-moving startup ecosystem, these methods are often insufficient. Artificial intelligence now introduces a shift by enabling continuous, data-driven risk monitoring and early anomaly detection.

AI systems, particularly those based on unsupervised learning and anomaly detection algorithms, are capable of flagging irregular patterns in startup behavior or performance long before they become critical, making it a necessary tool for VCs. With input as structured and unstructured data across multiple dimensions financial metrics, operational KPIs, user behavior, market sentiment, legal disclosures, and more, AI models can:

- **Identify unusual financial trends**, such as sudden drops in revenue, inflated valuations without correlated traction, or unexplained changes in burn rate
- **Monitor behavioral anomalies**, like abrupt changes in founder or executive activity, reduced engagement on public platforms, or inconsistent communications with investors
- **Detect external risk signals**, including lawsuits, regulatory flags, patent disputes, or significant shifts in customer reviews or media coverage
- **Score portfolio companies by risk level**, dynamically updating risk profiles and surfacing ventures that may require immediate attention or corrective measures

Advanced systems can even cross-reference internal data with public and third-party sources (Crunchbase, PitchBook, Glassdoor, Twitter, SEC filings) to produce a real-time risk landscape.

This level of insight enables VCs to:

- Intervene earlier in struggling companies, increasing the chances of course correction
- Allocate support resources more effectively, prioritizing high-risk ventures
- Adjust follow-on investment strategies based on emerging risk signals

- Communicate more transparently with LPs by backing decisions with real-time data.

Some venture capital firms are now integrating AI-driven risk engines directly into their internal evaluating systems. For example, Tribe Capital and EQT Ventures both use proprietary scoring systems that incorporate both opportunity and risk dynamics. Others leverage external platforms such as Zebra, Dealroom.ai, or Synaptic to augment their internal models.

Of course, risk models have limits. They could struggle with data sparsity in early-stage startups or misinterpret context-specific events as negative outliers. Therefore, AI should be seen not as a replacement for human risk managers, but as a powerful supporting tool to improve foresight and responsiveness and robustness.

In summary, AI-based risk management introduces a level of predictive vigilance previously unlocked to venture capital firms. By detecting early signals of distress and dynamically adjusting risk profiles, these tools upgrade the fund's ability to protect their capital, manage uncertainty, and make more informed, resilient decisions (ref. 22- Ramachandran, *Advanced Artificial Intelligence in Private Equity and Venture Capital: A Functional Framework for Lifecycle Transformation*, IBM Research, 2025).

4.6) Exit Forecasting and Return Estimation

Forecast of the exit is a critical component of venture capital strategy, as the ROI is realized mainly through events such as acquisitions, IPOs, or secondary sales. Predicting the probability, the timing, and the magnitude of these exits has traditionally been based on pattern recognition, comparable transactions, and experience-based intuition. Artificial intelligence now provides a structured and scalable alternative (once again), allowing firms to make more informed, based on data, exit strategies. AI models can now analyze historical exit data from thousands of startups and market conditions to identify patterns that link with successful outcomes. These models incorporate variables such as sector, funding stage, capital raised, market size, founder background, revenue growth, customer churn, investor reputation, and competitive dynamics. Through machine learning techniques, the system can then estimate the probability of an exit, its potential valuation, and the expected time frame. Predictive tools for exit estimation enable venture capital firms to:

- Prioritize follow-on investments based on updated return potential and exit forecasts
- Adjust portfolio strategies in response to shifting market conditions or macroeconomic trends
- Simulate exit scenarios under different assumptions and generate risk-adjusted return models
- Provide more accurate expectations to LPs regarding fund performance and capital recycling

Advanced models such as gradient boosting machines, survival analysis algorithms, and time-to-event prediction networks have shown promise in forecasting exits. Some funds have integrated these tools into their internal set of tools, while others use third-party platforms that offer exit scoring capabilities (ref. 23-Calafiore et al., *A Classifiers Voting Model for Exit Prediction of Privately Held Companies*, arXiv, 2019).

Additionally, return estimation models can be linked to dynamic startup performance. For example, if a company surpasses a key milestone such as 10x revenue growth in a 12-month period or expands into new markets, the model recalibrates projected outcomes. This adaptive capability ensures that the VC's internal financial projections remain current and actionable. However, AI-driven exit forecasting is still limited by data quality, outlier behavior, and the unpredictability of macro shocks such as financial crises or regulatory disruptions. For this reason, these models are best used in combination with expert judgment and qualitative insights.

4.7) Automated Benchmarking Between Similar Startups

Benchmarking is a fundamental step in venture capital that enables investors to assess a startup's progress, potential, and performance relative to peers in similar industries, stages, or locations. Usually, this comparison relies on manually compiled data, human analyst experience, and access to own deal flow insights. However, this approach, along with the others cited above, is time-consuming, inconsistent, and often limited by cognitive and informational bias. Artificial intelligence introduces a scalable and systematic alternative through automated benchmarking systems.

AI-based benchmarking tools work by analyzing structured and unstructured data from thousands of startup cases. These systems gather and standardize inputs such as funding history, revenue growth, user metrics, product adoption, hiring trends, burn rate, and digital engagement. Using techniques like clustering algorithms, semantic similarity analysis, and vector-based search, AI can identify startups that share comparable characteristics and performance trajectories.

This process allows venture capital firms to:

- Contextualize the performance of portfolio companies by comparing them to real-time peer groups
- Detect whether a startup is underperforming or outperforming relative to industry benchmarks
- Adjust valuation expectations and investment terms based on market comparables
- Identify market trends and emerging patterns across clusters of similar ventures

For example, if a startup in Series A stage shows 80% year-over-year growth but operates in a sector where the top quartile grows at 120%, AI tools can catch this gap and inform follow-up discussions. In a similar way, benchmarking instruments can surface lesser-known but high-performing startups that exhibit patterns like previous unicorns, enabling earlier and more confident investment decisions.

These tools are particularly powerful when integrated into the data ecosystem of the VC. By connecting benchmarking data with CRM data, KPI dashboards, and exit forecasts, venture capitalists can build a realist view of each investment's relative and absolute positioning. Platforms such as Harmonic, Zint, Dealroom, and custom-built internal solutions are examples of systems that provide these capabilities.

Nevertheless, automated benchmarking has its constraints. The definition of “similarity” is inherently subjective and varies by investment thesis, sector dynamics, and geographic focus. Additionally, the availability of reliable private company data can still be a limiting factor, especially for early-stage or stealth-mode startups (ref. 24-Petersone et al., *A Data-Driven Framework for Identifying Investment Opportunities in Private Equity*, arXiv, 2022).

Chapter 5

Data-driven VC case studies

After exploring theoretical frameworks and technical applications of artificial intelligence in previous chapters, it is now time to move from theory to practice, with some cases. While many venture capital firms are still experimenting with small-scale AI tests, a few pioneering funds and firms have already fully integrated data-driven systems into their internal operations. These firms are a testimony how AI can become a real strategic asset, rather than just an experimental support tool.

By looking at concrete case studies, it is possible to understand not only the potential of these technologies, but also the practical challenges and competitive advantages they create. Each example provides insights into how AI models, big data infrastructures, and predictive algorithms can transform daily work, from identifying promising startups to managing portfolio performance and supporting founders.

This chapter focuses on some of the most advanced firms and platforms that use and adopt AI extensively. Their cases demonstrate that combining data analytics with traditional investment skills can lead to faster decisions, broader market coverage, and ultimately better returns and outcomes for both general partners and limited partners.

5.1 SignalFire: Internal AI Engine for Scouting and Investment

SignalFire is one of the most advanced examples of a venture capital firm built around a data-first and AI-centric approach. Unlike traditional VCs, that depend mainly on personal connections, market rumors, and manual analysis, SignalFire has developed an internal technology platform capable of monitoring the startup ecosystem in real time and at scale.

At the center of SignalFire's approach is a proprietary AI engine that has the ability to continuously scan over two million data sources. These include job postings, GitHub activity, website traffic, app store rankings, and media mentions. For example, the system could notice that a particular startup has increased engineering hires by 30 percent in just a few months or that downloads of a new app have doubled in a short period. Such information, which are often invisible to the human analysts, allow SignalFire to identify emerging companies early, sometimes even before a formal fundraising round begins, rsulting in a strong competitive advantage.

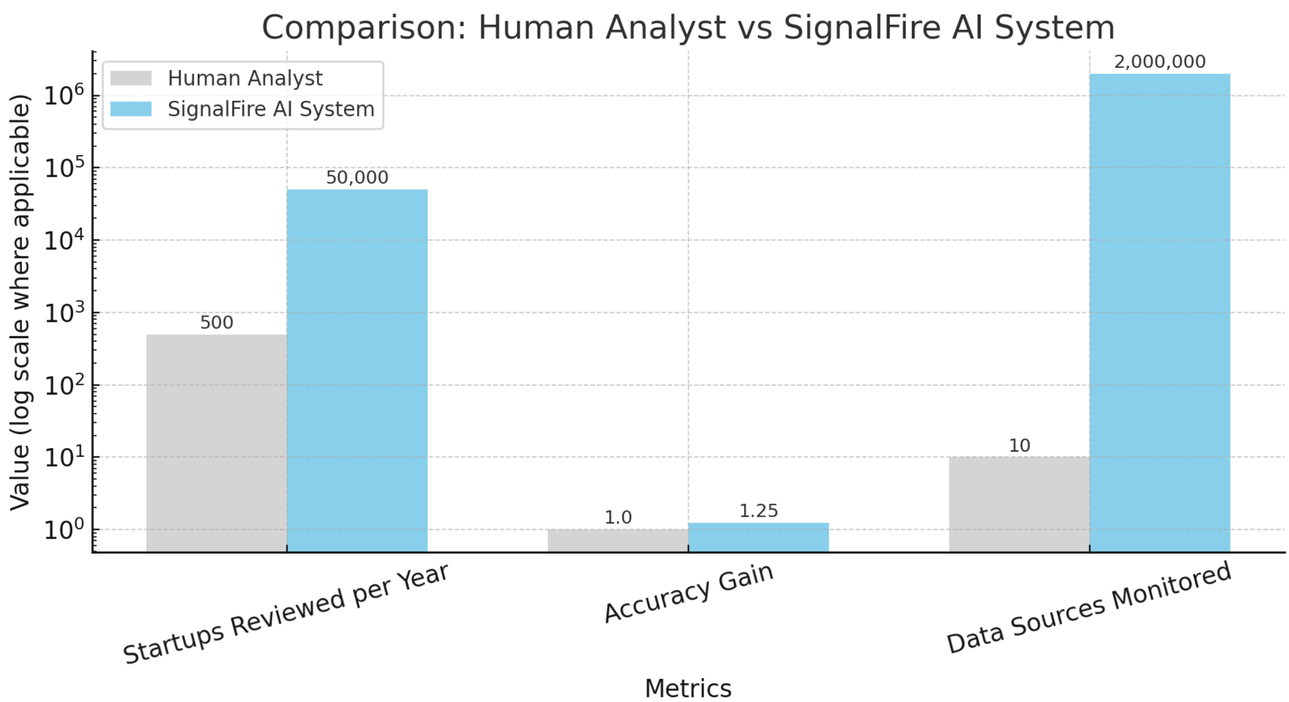
From a quantitative perspective, a traditional analyst might manually screen around 500 startups per year. In comparison, SignalFire's system can process information on more than 50,000 startups annually, increasing the reach and reducing the likelihood of missing promising opportunities. According to internal evaluations, this technological advantage improves the chance of identifying future high-performing startups by approximately 20 to 25 percent.

Another key element of SignalFire's model is its predictive abilities. The AI system not only monitors current activity but also estimates the probability that a startup will succeed in securing future funding rounds or achieving a successful exit. By analyzing variables such as user growth, customer churn, burn rate, and market dynamics, SignalFire can simulate potential scenarios and support investment decisions with quantitative evidence.

For instance, if a startup's data suggests a high probability (let us say, 85 percent) of raising a Series A round in a timeframe of 6 months, SignalFire may decide to engage early, offer bridge financing, or position itself as a lead investor. This proactive approach helps the fund stay ahead of competitors and establish stronger relationships with founders.

SignalFire also uses its AI infrastructure to support portfolio companies after the initial investment. By sharing market trends, talent movement insights, and operational benchmarks, the firm helps founders make better strategic decisions and anticipate challenges.

This combination of advanced data analytics and proactive support illustrates how AI can transform venture capital from a largely intuition-based business into a systematic and scalable strategy. SignalFire shows that when technology is deeply integrated into all phases of the investment process, it is possible to achieve faster, more informed, and more consistent decisions, ultimately creating a competitive edge in a highly crowded market (ref. 25-Signature Block, The Data Revolution in Venture Capital).



5.2 Inovia Capital: Use of Machine Learning for Predictive Decision-Making

Inovia Capital is a leading North American venture capital firm that is known for its data-driven investment approach and strong focus on long-term partnership with founders.

Unlike firms that rely mainly on traditional financial metrics and intuition, Inovia has integrated machine learning into several steps of its functional processes, making a more structured and evidence-based evaluation of startups.

At the centre of Inovia's strategy is the adoption of predictive models to determine the growth potential and financial sustainability of startups. These models analyze a wide range of data points, including monthly recurring revenue trends, churn rates, customer acquisition costs, and lifetime value estimates. By applying machine learning algorithms to this data, Inovia can estimate the future revenue trajectories and can identify early warning signals that might indicate operational failures or risks.

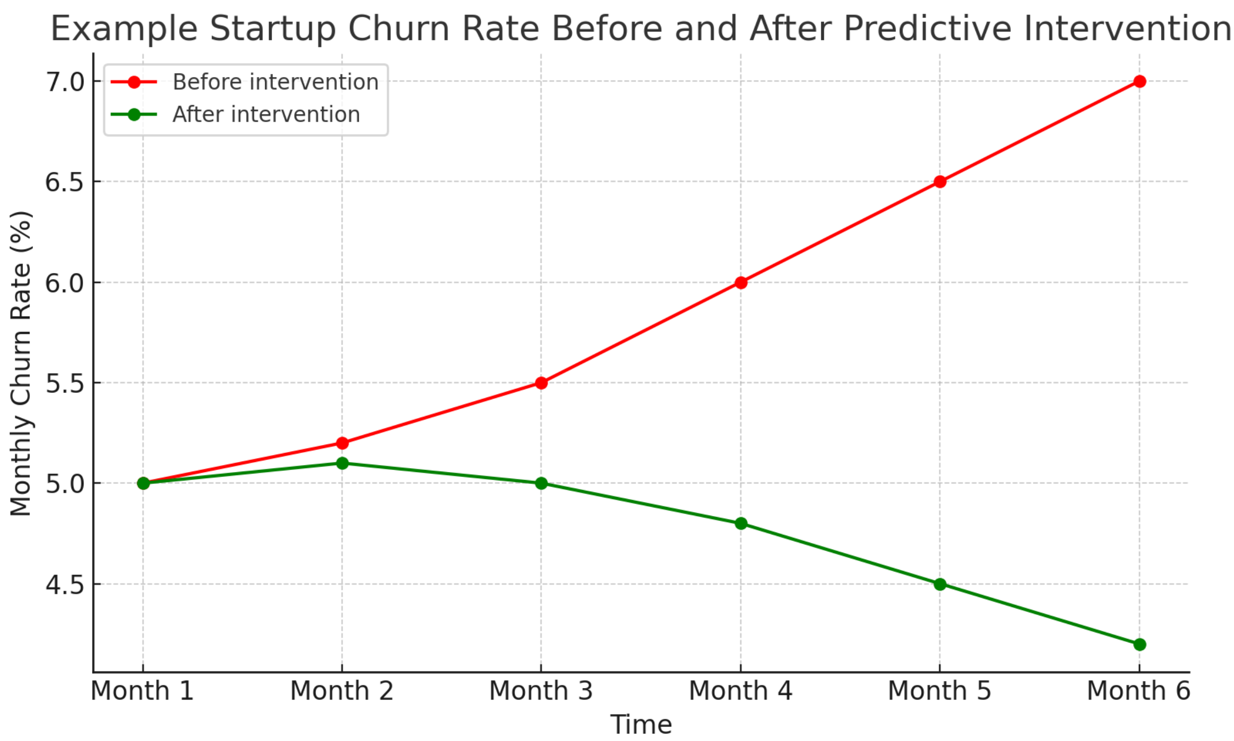
For example, in one of its early-stage portfolio companies, Inovia used predictive models to detect a subtle but consistent increase in customer churn that was not immediately visible (at naked eye) in headline growth figures. This insight allowed the investment team to join and work with the founders on improving retention strategies before the problem occurred, ultimately helping the company stabilize and secure a successful Series B round.

Quantitatively, Inovia's machine learning tools process data on approximately 5,000 startups per year, compared to the few hundred that a traditional analyst might evaluate in detail. So, these models can simulate different market scenarios and test possible financial forecasts, providing a probabilistic view of future performance. According to internal reports, this method has improved the accuracy of growth predictions by about 15 to 20 percent and reduced due diligence time by nearly 30 percent, which has huge value.

In addition to growth forecasting, Inovia also uses natural language processing techniques to analyze founder communications, investor updates, and market sentiment data. By automatically classifying and summarizing large volumes of text, the firm gains a more nuanced understanding of a startup's strategic positioning and potential risks.

Inovia's use of machine learning demonstrates that advanced analytics and AI models can coexist with human judgment, rather than replace it, a recursive theme across this study. The final investment decision still relies on the experience and intuition of partners, it will always also be their responsibility, but these decisions are now supported by quantitative insights and data and a more comprehensive understanding of each startup's dynamics. This hybrid approach allows Inovia to identify opportunities earlier, act more confidently, and provide better support to its portfolio companies throughout their journey.

In summary, Inovia Capital's integration of machine learning into its investment process illustrates how data science can enhance both the speed and the quality of venture capital decisions, while still valuing the human element at the core of successful investing (ref. 26- Inovia Capital, Machine Learning and Predictive Analytics in Venture Capital).



5.3 Tribe Capital: "Quantitative VC" Approach and Proprietary Scoring System

Tribe Capital is a Silicon Valley-based venture capital firm that has become widely known for its “quantitative VC” approach. Rather than relying mainly on intuition or human signals, Tribe integrates deep data analysis and proprietary algorithms into almost every stage of its investment process. This strategy aims to bring a more scientific and reproducible method to early-stage investing, traditionally seen as highly subjective.

The main advantage given by Tribe’s approach is its proprietary scoring framework, known as the "Data-Driven Product-Market Fit Signal." This system is created to measure how strongly a startup’s product connects with its users, which is considered one of the most critical indicators of future success. The framework uses direct data from a startup's internal systems, including user retention curves, daily and monthly active user ratios (DAU/MAU), engagement depth, and referral metrics.

For example, a startup with 30 percent month-over-month user growth and a DAU/MAU ratio above 50 percent is more likely to receive a much higher score than one with slower growth and lower engagement, even if the latter has more revenue at an early stage. By analyzing these usage patterns, Tribe can estimate not only current traction but also predict future growth potential with greater confidence.

Numerically, Tribe’s team analyzes data from over 1,000 startups annually, using machine learning models to simulate different scenarios and benchmark each company against a historical database of both successful and failed startups. According to internal analysis, this method has improved the firm’s ability to identify high-performing startups by approximately 25 percent compared to a purely qualitative approach.

Beyond product-market fit, Tribe also incorporates data on the founding team, social graph connectivity, and network effects. For example, by mapping the relationships between founders and their past collaborations or investors, the firm can assess the strength and resilience of the team in facing market challenges. Even more, by tracking fundraising activity and external signals and flags like press mentions or job postings, Tribe can dynamically update its risk and opportunity scores.

Once again, Tribe Capital's philosophy demonstrates that quantitative methods do not need to replace human judgment. Partners and analysts still meet founders, evaluate strategic vision, and discuss long-term plans, but these subjective impressions are anchored in robust, data-driven assessments. This dual approach allows Tribe to act quickly when it detects strong signals and to avoid deals where metrics suggest underlying weaknesses.

Tribe Capital's quantitative investment model represents an evolution of the traditional VC approach, showing that integrating data science into decision-making can improve accuracy, reduce bias, and increase the chances of backing startups with true long-term potential (ref. 27-TechCrunch, VC Arjun Sethi talks about Tribe Capital's quantitative strategies and Termina platform, TechCrunch).

5.4 EQT Ventures and the "Motherbrain" Tool

EQT Ventures is a European venture capital firm that stands out for its heavy investment in proprietary technology to support decision-making. They created their own AI platform, known as "Motherbrain," which plays a central role in the firm's scouting and evaluation processes. EQT Ventures leverages Motherbrain to continuously monitor and evaluate thousands of startups globally.

Motherbrain is designed to analyze large amounts of structured and unstructured data, including web traffic metrics, app downloads, employee growth trends, social media activity, and press coverage. For example, if a startup shows a sudden increase in organic web traffic or app engagement, Motherbrain flags it for further analysis. This helps EQT Ventures identify potential investment opportunities that might be overlooked through traditional sourcing methods.

Numerically, Motherbrain tracks data on more than 1 million startups globally. According to EQT Ventures, around 35 percent of their portfolio companies have been sourced or initially identified through insights provided by Motherbrain.

One practical example is investment in Handshake, a career platform for students and recent graduates. Motherbrain detected unusual spikes in user engagement and employer sign-ups, prompting the investment team to investigate further, ultimately leading to a successful funding round.

In addition to scouting, Motherbrain is used to support due diligence and post-investment monitoring. By aggregating real-time performance metrics and comparing them to peer benchmarks, the tool helps partners make more informed decisions about the follow up of the investments, operational interventions, or potential exit strategies.

Another strength of Motherbrain lies in its ability to reduce human bias. Traditional sourcing methods often favor startups in certain geographies or those led by founders within established networks. Motherbrain allows EQT to take a more data-driven, objective approach, potentially uncovering strong teams and products that would otherwise remain under the radar.

Overall, EQT Ventures' investment in Motherbrain demonstrates how an AI-first approach can transform venture capital from a largely relationship-driven business into a scalable and systematic process. By combining advanced analytics with human expertise, EQT Ventures goal is to improve decision quality, enhance portfolio performance, and increase access to innovative founders across markets (ref. 28-EQT Ventures, Motherbrain Platform Overview).

5.5 GV (Google Ventures): Internal AI and Analytical Tools

GV, formerly known as Google Ventures, is the venture capital of Alphabet and one of the main pioneers when it comes to integrating new technologies and AI into venture investing. Again, unlike many traditional VC funds that depend mainly on qualitative assessments and partner intuition, GV leverages Google's deep expertise in data science to support a more structured and scalable investment process.

GV has built a sophisticated internal data analytics team that works alongside investment partners to analyze potential deals. This team uses a wide range of machine learning models to evaluate market trends, simulate growth scenarios, and assess operational metrics of startups. One core asset of GV’s approach is its ability to combine internal startup data with external market signals, such as search trends, online engagement metrics, and broader industry indicators.

For example, GV often uses proprietary models to analyze web traffic and search query patterns as early indicators of market demand or user traction. If a consumer product startup shows a sudden spike in organic search interest, this can signal strong potential demand, prompting the investment team to prioritize further due diligence.

Numerically, GV’s data analytics infrastructure makes it possible for the team to evaluate thousands of startups each year, far beyond what a traditional analyst team could manually cover. This expanded coverage increases the likelihood of identifying breakout companies early. According to internal reports, integrating advanced analytics has shortened the average time to reach a term sheet decision by around 25 percent and improved the accuracy of revenue growth forecasts by 15 to 20 percent.

Beyond sourcing and screening, GV applies AI and analytics to post-investment monitoring. Portfolio companies are regularly benchmarked against anonymized data from similar ventures, enabling founders and GV partners to identify strengths and weaknesses in real time. This continuous monitoring helps guide strategic decisions, such as when to scale operations or enter new markets.

GV also emphasizes support beyond capital. By providing startups with access to Alphabet’s technical resources and data expertise, GV helps its portfolio companies make more informed product and market decisions. This partnership approach combines the strengths of human judgment, industry knowledge, and cutting-edge analytical tools.

In summary, GV's integration of internal AI and data science illustrates how large-scale analytics can transform venture capital from a network-driven business into a data-augmented, evidence-based discipline. This hybrid approach enhances decision-making speed and accuracy while still leaving room for the human insight that remains essential to successful investing (ref. 29-GV, Our Data-Driven Approach, GV Insights).

Chapter 6

Limits, challenges and future perspectives

While the use of Artificial Intelligence into Venture Capital processes offers significant advantages for efficiency, scalability and better decision-making, it is also crucial to recognize its current limitations and the challenges that accompany its widespread adoption.

AI models are powerful analytical tools, but they can fail if misused. They depend heavily on the quality and completeness of the data they are trained on, and they often operate as “black boxes,” making it difficult for investors and users to fully understand or trust their outputs. Furthermore, the use of AI introduces new ethical questions related to bias, transparency, and fairness in startup selection, which can affect both the market and society.

From a practical point of view, there is also a risk of excessive automation that could overwhelm the fundamental human elements of venture investing, such as intuition, gut feeling, experience, and relationship-building. The risk is that investors might rely too blindly on models, potentially missing unique qualitative signals that only human judgment can capture, this risk also affects startup’ funders that show promise but go unnoticed.

At the same time, the rapid development of AI technologies opens fascinating scenarios for the future of venture capital, one of them is the emergence of hybrid models, that work by combining human expertise with advanced algorithms, or even the possibility of fully automated investment funds.

This chapter aims to explore these technical and ethical challenges in detail, analyze the risks connected to over-reliance on automation, and provide a forward-looking discussion on how venture capital might evolve in the coming years, shaped by continuous AI innovation and changing market dynamics.

6.1 Technical limitations: Opacity, incomplete data, and generalizability

Even if AI has brought significant advances and upgrades to venture capital industry, it still has several technical limitations that block its full potential. As cited above, one of them is in the so-called "black box" nature of most AI models. In other words, these systems can make predictions or suggestions without clearly explaining how they arrived at those results, both for users and developers. Comprehensively, for many investors, this can create hesitation. After all, when large amounts of money are at stake, being unable to fully understand or justify an algorithm's recommendation can feel risky and even irresponsible (ref. 30-Doshi-Velez & Kim, Towards a rigorous science of interpretable machine learning, arXiv, 2017).

Another major challenge involves the quality and completeness of data used to train these algorithms. Unlike sectors such as e-commerce or online advertising early-stage startups often lack reliable and detailed information. Many young companies don't have long track records, consistent financial statements, or standardized market metrics. As a result, AI models can only analyze what they have access to, potentially missing subtle but critical factors that an experienced investor might pick up in a conversation with the founder or during a site visit.

A third limitation is that AI models trained in a specific environment may not work well in different contexts — a problem known as generalizability. For example, a model fine-tuned on U.S. startup data might struggle to evaluate a new business in Southeast Asia or Africa, where market conditions, consumer behaviors, and regulatory frameworks can be completely different. This lack of adaptability can lead to false positives or overlook promising ventures that don't "fit" the patterns the model was trained on (ref. 31-Bengio et al., Deep learning for AI, Communications of the ACM, 2021).

These technical weaknesses highlight the importance of keeping human judgment at the center of investment decisions. While algorithms can sift through data quickly and spot patterns humans might miss, they are not yet capable of replacing the nuanced evaluations that experienced partners provide. Instead, a hybrid approach — where AI acts as an intelligent assistant rather than a decision-maker — appears to be the most reliable path forward.

By being aware of these limitations, VC firms can better calibrate their use of AI, avoiding over-reliance on models and preserving the creative, intuitive side of venture investing. Ultimately, while AI offers impressive tools, it is the combination of data-driven insights and human intuition that defines the best outcomes in such a complex and unpredictable sector.

6.2 Ethical issues: Model bias and transparency toward startups

Beyond technical constraints, the integration of AI into venture capital introduces important ethical challenges, particularly concerning model bias and the need for transparency in interactions with startups.

AI systems learn from historical data, which often contain embedded biases reflecting past decisions and social inequalities. For example, if historical funding data favored certain founder profiles (such as male or Silicon Valley-based teams), the models trained on these datasets might replicate or even amplify these tendencies. This raises serious concerns regarding fairness and equal opportunity for entrepreneurs from diverse backgrounds or underrepresented regions.

Bias can manifest at multiple levels: from the initial filtering of startups to risk scoring, to final investment recommendations. The danger is that AI-driven decisions may appear objective and neutral, when in fact they reinforce systemic barriers that the venture capital industry has struggled to overcome for decades.

In addition to bias, transparency is a crucial ethical consideration. Startups interacting with VCs deserve to understand how and why certain investment decisions are made, especially when those decisions are heavily influenced by algorithms. However, the opacity of most advanced AI systems complicates the communication of decision rationales. As a result, founders may feel excluded from processes or unfairly evaluated, damaging trust and potentially discouraging promising projects from seeking funding (ref. 32-Liu, Discrimination in the VC funding process: The role of algorithms, arXiv, 2020).

Some firms have attempted to address these ethical concerns through explainable AI (XAI) techniques, which aim to make algorithmic outputs more interpretable and understandable. While XAI is a promising step, it remains technically complex and often insufficient to fully mitigate deeper biases present in the data or design.

Overall, the ethical risks posed by AI in venture capital highlight the importance of combining technological adoption with a strong governance framework. Regular bias audits, diverse data sources, and proactive disclosure policies can help ensure that AI systems support rather than undermine fairness and inclusivity in startup funding.

6.3 Risks of excessive automation in decision-making processes

While AI offers significant efficiency and scalability advantages, relying too heavily on automated decision-making in venture capital can introduce serious strategic and operational risks.

The first major concern is the potential loss of human intuition and context-specific judgment. Successful venture investing often depends on subtle, qualitative signals that go beyond quantitative data: a founder's charisma, the uniqueness of a vision, or informal market insights gathered through personal interactions. Fully automated models, even when highly sophisticated, are not capable of integrating these uniquely human dimensions, which are crucial in identifying truly exceptional opportunities (ref. 33-McKinsey & Company, The state of AI in 2021, McKinsey Global Survey, 2021).

A second risk is the creation of algorithmic herd behavior. When many venture firms use similar data sources and models, investment decisions may converge toward the same types of startups, reducing diversity and inflating valuations artificially. This concentration can lead to speculative bubbles and destabilize markets.

Additionally, excessive automation can weaken the internal dynamics of VC firms. Over-reliance on algorithms might reduce internal debate, creative thinking, and collective expertise, transforming partners and analysts into mere overseers rather than active evaluators. This shift can harm organizational culture and diminish the innovative spirit needed to identify and support groundbreaking ventures.

Finally, when decision-making is heavily automated, it becomes more difficult to assign responsibility for failed investments or ethical lapses. In traditional setups, partners are directly accountable to limited partners and stakeholders. With algorithmic decisions, accountability is more diffuse, creating new legal and governance risks (ref. 34-Crane, Data science and the art of herd detection, arXiv, 2020).

And so, while AI can improve operational efficiency and enhance screening, it is crucial for venture capital firms to maintain a balance between automated insights and human judgment. Hybrid approaches, where AI serves as an augmentation rather than a replacement, represent a more robust and responsible path forward.

6.4 Impact on human labor in the sector (Traditional VCs vs. Data Scientists)

The adoption of AI in venture capital is not only transforming processes and decision-making but is also reshaping the roles and skills required within firms. Traditionally, VC teams have relied heavily on human expertise, interpersonal skills, and intuition to evaluate startups, build relationships, and guide founders through growth phases.

As data-driven approaches gain traction, the demand for technical roles such as data scientists, machine learning engineers, and AI strategists is rapidly increasing. This shift is creating a new hybrid workforce where traditional investment professionals work alongside technical experts to analyze complex data and design algorithmic models (ref. 35- Gompers, How do venture capitalists make decisions? Journal of Financial Economics, 2020).

However, this transition raises questions about the future relevance of traditional VC skills. While relationship management and qualitative assessment remain crucial, there is a risk that human judgment could be overshadowed by algorithmic outputs, reducing the influence of experienced partners in strategic decisions. Some firms might prioritize quantitative signals over soft factors like founder motivation or cultural fit, potentially altering the DNA of venture investing.

Additionally, the cultural shift toward more analytical and technical decision-making can create tension within VC organizations. Investment professionals may feel challenged by the growing influence of technical staff, leading to internal friction or resistance to new methods. At the same time, data scientists might struggle to fully understand or value the nuanced, relationship-based aspects of startup evaluation, highlighting a potential skills gap between technical and traditional investment perspectives.

Ultimately, the integration of AI requires a balanced and collaborative approach where both technical and human-centered expertise are valued. The ability to bridge these two worlds will likely define the success of future venture capital firms, enabling them to remain competitive while preserving the human elements that are essential to supporting founders and building lasting partnerships.

6.5 Future perspectives: Hybrid models, capital democratization, fully automated funds

Looking ahead, the integration of AI in venture capital is expected to transform not only operational processes but also the fundamental logic behind investment strategies. One promising development is the rise of hybrid models, where AI algorithms and human expertise work together in a complementary way. In this scenario, AI is used to analyze large volumes of market and startup data, prioritize deals, and generate quantitative risk assessments, while human partners focus on interpreting nuanced signals, mentoring founders, and building long-term strategic visions (ref. 36-Stanford HAI, Artificial Intelligence Index Report 2023).

Another potential trajectory is capital democratization. By providing scalable analytical tools and improving transparency, AI-powered platforms may lower entry barriers and allow more diverse investors to participate in early-stage funding. This can help reduce the concentration of capital among a few elite VC firms and promote more equitable access to resources for startups worldwide.

Finally, there is speculation about the emergence of fully automated VC funds, in which algorithms independently manage all phases of investment, from sourcing and due diligence to exit strategies. While this idea illustrates the extreme potential of AI, it faces serious obstacles: lack of transparency, difficulty in capturing qualitative insights, ethical challenges, and reduced trust from founders and LPs. Current evidence suggests that while partial automation is likely to grow, human judgment will remain central in high-stakes investment decisions (ref. 37-De Cremer & Kasparov, AI should augment human intelligence, not replace it, Harvard Business Review, 2021).

Rather than a complete replacement of human expertise, the future of venture capital will likely be shaped by thoughtful integration, where firms blend AI-driven scalability with the relational and strategic depth provided by experienced partners. Those who succeed in finding this balance will define the next generation of venture capital leaders.

Chapter 7

Conclusion

This final chapter brings together the main findings and reflections that have emerged throughout the thesis, offering a comprehensive synthesis of the work. After analyzing the structure and logic of venture capital, exploring the technological foundations of AI, and discussing practical applications and limitations, it is now time to draw conclusions on how these elements interact and what they imply for the future of the sector.

The chapter begins by summarizing the key results derived from research and case studies, highlighting both the potential and the constraints of integrating AI into venture capital processes. It then offers personal considerations on the journey of this study, including critical insights and reflections on the practical implications for investors and entrepreneurs. Finally, it closes with overall considerations, outlining possible future directions and recommendations for further research.

By concluding this work, the intention is not only to provide a clear overview of what has been discovered, but also to stimulate deeper reflection on the evolving relationship between human expertise and technological innovation in high-risk investment contexts.

7.1 Results

The analysis developed throughout this thesis confirms that the integration of artificial intelligence into venture capital is not a passing trend or purely theoretical concept, but rather a deep structural evolution already underway in many leading firms worldwide. By systematically examining AI technologies, practical applications, case studies, and ethical challenges, it becomes clear that data-driven approaches are redefining how VCs operate, select startups, and generate value.

One of the main findings is the concrete contribution of AI to deal with sourcing and scouting activities, traditionally one of the most resource-intensive phases for VC firms. Thanks to the ability to process vast amounts of data from multiple sources — including patent databases, social media, founder networks, and market analytics — AI models can identify early signals of promising startups long before they appear on the radar of traditional investors. This capability not only accelerates the discovery process but also improves accuracy by reducing the influence of human biases and subjective impressions.

Further key result concern due diligence and evaluation processes, which benefit significantly from advanced analytical tools. Predictive models trained on historical success and failure data can suggest probability scores for startup survival, growth potential, or likelihood of future funding rounds. Natural language processing tools can analyze founder pitches, interview transcripts, and online content to detect sentiment, consistency, and credibility, offering a more comprehensive and objective complement to traditional qualitative assessments.

Case studies discussed in this thesis, including Signal Fire and its internal AI engine, Tribe Capital with its proprietary "quantitative VC" scoring system, and EQT Ventures'

Motherbrain platform, highlight real-world evidence of these advantages. These firms have demonstrated improvements not only in the speed of evaluation but also in overall investment performance, measured by higher success rates in portfolio companies and more efficient resource allocation. Their experiences underline the practical impact of AI when effectively integrated into decision-making workflows.

Another important outcome relates to portfolio monitoring and risk management. The ability of AI systems to track large volumes of operational data in real time enables proactive identification of underperformance signals, potential cash flow problems, or market misalignments. By providing early alerts, the resources or VCs to intervene strategically, support startups with targeted resources, or adjust investment strategies to mitigate potential losses. This dynamic, data-driven approach contrasts sharply with traditional, static monitoring practices based on periodic reports and manual updates.

However, the findings also reveal significant technical and ethical constraints that limit the full automation of venture capital processes. The opacity of many AI models, often referred to as the "black box" problem, remains a central barrier to broader adoption. Investors need clear and explainable rationales to justify decisions to limited partners and other stakeholders, and purely algorithmic outputs can undermine trust if they are not transparent or interpretable.

Furthermore, the issue of data bias emerges as a critical challenge. Historical funding data often reflect existing societal inequalities and preferential patterns (such as favoring male founders or specific geographic hubs like Silicon Valley). Without deliberate correction mechanisms, AI models trained on these datasets risk perpetuating and even amplifying these biases, potentially reinforcing systemic barriers for underrepresented founders. This insight underscores the importance of combining technical rigor with ethical governance when deploying AI in venture contexts.

Another result of the study concerns the impact on human roles and skills within VC firms. As AI tools take over analytical and repetitive tasks, traditional investment professionals are called to evolve their competencies. Rather than being replaced, their roles shift toward higher-level strategic thinking, relationship management, and nuanced judgment, areas where human intuition and experience remain irreplaceable. The firms that successfully foster collaboration between data scientists and investment partners seem best positioned to leverage the complementary strengths of both human and machine intelligence.

Finally, the exploration of possible future scenarios suggests that the most likely path forward lies in hybrid models, where AI augments but does not replace human decision-making. While experiments with fully automated funds are technologically intriguing, they face considerable resistance due to accountability concerns, legal uncertainties, and the irreplaceable value of human insight in high-stakes, context-sensitive investment decisions. Moreover, the potential for AI to democratize access to venture capital, making sophisticated analysis tools available to smaller investors and founders, represents an exciting opportunity, but it also requires robust safeguards to prevent speculative bubbles and ensure market stability.

Overall, the results of this thesis highlight that AI can significantly improve operational efficiency, decision quality, and market reach for VC firms. However, these advantages can only be fully realized through thoughtful and responsible integration strategies that respect both technical limitations and ethical imperatives. Combining the scale and analytical power of AI with the relational, creative, and strategic capabilities of human investors appears to be the most promising approach for the future of venture capital.

7.2 Personal considerations

As an engineer, I firmly believe in progress and innovation, and I am convinced that artificial intelligence represents a true turning point in the evolution of the human species. It is understandable to feel skeptical or distrustful of something we do not fully comprehend; however, it is precisely those who choose to embrace it who will reap the greatest benefits.

During the "Internet revolution," the world was dramatically reshaped: companies that invested early became global leaders, and individuals who understood its potential achieved remarkable success and influence. Today, we are witnessing a similar transformation with AI.

In our current world, venture capitalists play a crucial role in identifying and supporting the future leaders, both companies and individuals. AI is set to further enhance this selection process, making it faster, more accurate, and ultimately more impactful.

AI is undoubtedly a powerful tool, and it is our responsibility to make a conscious effort to use it wisely and ethically. Personally, I made use of AI during the drafting of this thesis, and I must admit it significantly improved both time efficiency and the overall quality of my work.

Looking ahead, I can only imagine what AI will be capable of achieving for future generations. Hopefully, it will contribute to a world where they can enjoy an even better quality of life than I do today.

7.3 overall considerations

Looking at the analysis, it emerges that the introduction of AI into the venture capital sector is not a simple technological upgrade but a fundamental change in mindset. Each chapter has shown how AI can intervene at different stages: from improving deal flow efficiency and team evaluations to supporting strategic decisions and exit forecasts. The diverse applications suggest that AI is becoming a transversal enabler rather than a niche tool.

The examination of real cases like Tribe Capital and EQT Ventures has confirmed that data-driven approaches are already reshaping internal processes and competitive positioning. At the same time, the discussion on technical and ethical limits has highlighted that these innovations must be adopted with awareness and critical sense.

Overall, the integration of AI into VC appears as a gradual but irreversible path. It calls for a balance between analytical precision and entrepreneurial intuition, between automation and human value. This synergy could define the next generation of venture investing, combining data science with the core principles of value creation and long-term vision.

Acknowledgements

First of all, I would like to thank my thesis advisor, Professor Riccardo Calcagno, who has been quick and reliable, he gave me the opportunity to write my thesis on a topic I am truly interested in and kept an open mind, allowing me creative freedom in the way I developed and structured it.

Then, and most important, I want to thank my wonderful family, who have always supported and believed in me at every moment of my life. They are the one thing I allow myself to brag about. My love for them, and theirs for me, goes beyond the capability of human expression.

Moving forward, I want to thank the friends, starting with Lucio: my very friend and the greatest and most outstanding person, we shared the brightest and darkest moments of our academic path. Then, I want to thank Mido, Pippo, Albi, Viots, Samu, Edo, Mata, Febio, Giuseppe and Joe. They made this period full of joyful moments and have always encouraged me to push myself to the limits.

A special and unique thanks to my roommate Alessandro with whom I have a very engineering-like, constructive and very enjoyable relation, and to Catalina who always cheered for me.

References:

1- ref. 1 -Metrick & Yasuda, *Venture Capital and the Finance of Innovation*, Journal of Economic Literature, 2010 - https://academic.oup.com/rfs/article/23/6/2303/1582697
2- ref. 2- Kerr, Nanda & Lerner, *Entrepreneurship as Experimentation*, Journal of Economic Perspectives, 2014 - https://www.aeaweb.org/articles?id=10.1257/jep.28.3.25
3- ref. 3-Kaplan & Strömberg, *Financial Contracting Theory Meets the Real World*, NBER Working Paper, 2000 - https://www.nber.org/papers/w7660
4- ref. 4- Kerr & Nanda, *Financing Young Firms*, Journal of Finance, NBER Working Paper, 2009 - https://www.nber.org/papers/w14723

5- ref. 5-OECD, *Venture Capital Investments in Artificial Intelligence*, 2023 - https://www.oecd.org/publications/venture-capital-investments-in-artificial-intelligence-f97beae7-en.htm
6- Ref. 6-Chollet, F. *Deep Learning with Python,2021 -* https://livebook.manning.com/book/deep-learning-with-python-second-edition
7- (ref. 7-Xu & Tian, *A Comprehensive Survey of Clustering Algorithms*, Annals of Data Science, 2015 - https://link.springer.com/article/10.1007/s40745-015-0040-1
8- ref. 8- Bommasani et al., *On the Opportunities and Risks of Foundation Models*, Stanford CRFM, 2021 - https://arxiv.org/pdf/2108.07258
9- ref. 9-Burrell, *How the machine 'thinks': Understanding opacity in machine learning algorithms*, Big Data & Society, 2016 - https://journals.sagepub.com/doi/full/10.1177/2053951715622512
10- ref. 10 -Scabbio, *Exploring the Adoption of Artificial Intelligence in Venture Capital*, 2024 - https://research-api.cbs.dk/ws/portalfiles/portal/92143006/1518505_MT_Scabbio_148657.pdf
11- ref. 11-Ozince & Ihlamur, *Automating Venture Capital: Founder assessment using LLM-powered segmentation, feature engineering and automated labeling techniques*, arXiv, 2024 - https://arxiv.org/pdf/2407.04885
12- ref. 12-Montanaro et al., *Venture capital investments in artificial intelligence*, Journal of Evolutionary Economics, 2024 - https://link.springer.com/article/10.1007/s00191-024-00857-7
13- ref. 13-Fang, Tao & Li, *Anchoring AI Capabilities in Market Valuations: The Capability Realization Rate Model and Valuation Misalignment Risk*, arXiv, 2025 - https://arxiv.org/pdf/2505.10590
14- ref. 14-Potanin et al., *Startup success prediction and VC portfolio simulation using CrunchBase data*, arXiv, 2023 - https://arxiv.org/pdf/2309.15552
15- ref. 15-Cao et al., *Using Deep Learning to Find the Next Unicorn: A Practical Synthesis*, IJCAI FinNLP/Muffin Workshop, 2023 - https://arxiv.org/pdf/2210.14195
16- ref. 16-Calafiore et al., *Survival and Neural Models for Private Equity Exit Prediction*, arXiv, 2019 - https://arxiv.org/pdf/1911.08201
17- ref. 17-Ozince & Ihlamur, *Automating Venture Capital: Founder assessment using LLM-powered segmentation, feature engineering and automated labeling techniques*, arXiv, 2024 - https://arxiv.org/pdf/2407.04885
18- ref. 18-Tran et al., *Parsing Speech: A Neural Approach to Integrating Lexical and Acoustic-Prosodic Information*, arXiv, 2017 - https://arxiv.org/pdf/1704.07287
19- ref. 19-Arroyo et al., *Assessment of Machine Learning Performance for Decision Support in Venture Capital Investments*, IEEE Access, 2019 - https://doi.org/10.1109/ACCESS.2019.2938659
20- ref. 20-Lyu et al., *Graph Neural Network Based VC Investment Success Prediction*, arXiv, 2021 - https://arxiv.org/pdf/2105.11537
21- ref. 21-Bai & Zhao, *Analyzing the Effectiveness of Artificial Intelligence in Due Diligence in Mergers & Acquisitions*, 2025 - https://www.researchgate.net/publication/382951089_AI_in_Mergers_and_Acquisitions_Analyzing_the_Effectiveness_of_Artificial_Intelligence_in_Due_Diligence
22- ref. 22-Ramachandran, *Advanced Artificial Intelligence in Private Equity and Venture Capital: A Functional Framework for Lifecycle Transformation*, IBM Research, 2025 - https://www.researchgate.net/publication/390948235_Advanced_Artificial_Intelligence_in_P

rivate_Equity_and_Venture_Capital_A_Functional_Framework_for_Lifecycle_Transformation

23- ref. 23-Calafiore et al., *A Classifiers Voting Model for Exit Prediction of Privately Held Companies*, arXiv, 2019 - https://arxiv.org/pdf/1910.13969
24- ref. 24-Petersone et al., *A Data-Driven Framework for Identifying Investment Opportunities in Private Equity*, arXiv, 2022 - https://arxiv.org/pdf/2204.01852
25- ref. 25-Signature Block, The Data Revolution in Venture Capital - https://www.signatureblock.co/articles/the-data-revolution-in-venture-capital
26- ref. 26-Inovia Capital, Machine Learning and Predictive Analytics in Venture Capital - https://www.inovia.vc/perspectives/machine-learning-and-predictive-analytics-in-venture-capital
27- ref. 27-TechCrunch, VC Arjun Sethi talks about Tribe Capital's quantitative strategies and Termina platform, TechCrunch - https://techcrunch.com/2024/03/17/vc-arjun-sethi-talks-a-big-game-about-selling-his-company-picking-strategies-to-other-investors-he-says-theyre-buying-it/
28- ref. 28-EQT Ventures, Motherbrain Platform Overview - https://eqtventures.com/motherbrain
29- ref. 29-GV, Our Data-Driven Approach, GV Insights - https://www.gv.com/insights/our-data-driven-approach/
30- ref. 30-Doshi-Velez & Kim, Towards a rigorous science of interpretable machine learning, arXiv, 2017 - https://arxiv.org/pdf/1702.08608
31- ref. 31-Bengio, Deep learning for AI, Communications of the ACM, 2021 - https://cacm.acm.org/magazines/2021/7/253464-deep-learning-for-ai/fulltext
32- ref. 32-Liu, Discrimination in the VC funding process: The role of algorithms, arXiv, 2020 - https://arxiv.org/abs/2006.10935
33- ref. 33-McKinsey & Company, The state of AI in 2021, McKinsey Global Survey, 2021 - https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2021
34- ref. 34-Crane, Data science and the art of herd detection, arXiv, 2020 - https://arxiv.org/abs/2007.06570
35- ref. 35-Gompers, How do venture capitalists make decisions? Journal of Financial Economics, 2020 - https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2801385
36- ref. 36-Stanford HAI, Artificial Intelligence Index Report 2023 - https://aiindex.stanford.edu/report/
37- ref. 37-De Cremer & Kasparov, AI should augment human intelligence, not replace it, Harvard Business Review, 2021 - https://hbr.org/2021/03/ai-should-augment-human-intelligence-not-replace-it