



InData Labs

BIG DATA

in Banking and Finance

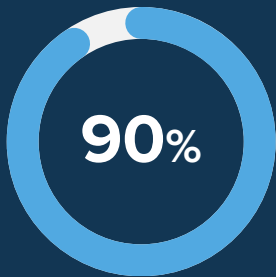


Table of contents

About Big data	3
Big data and generative AI: How they align and complement each other	7
Big data and AI in banking and finance	8
Banking and finance. Case studies	12
1. Fraud detection	12
2. Contact center efficiency optimization	13
3. Customer churn analysis	15
4. Risk management	16
5. Investment data management solution	17
Summary: Big data and AI use cases in finance	18
About InData Labs	19

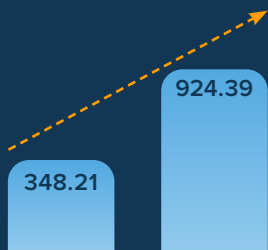
ABOUT BIG DATA

We create 2.5 quintillion bytes of data every day



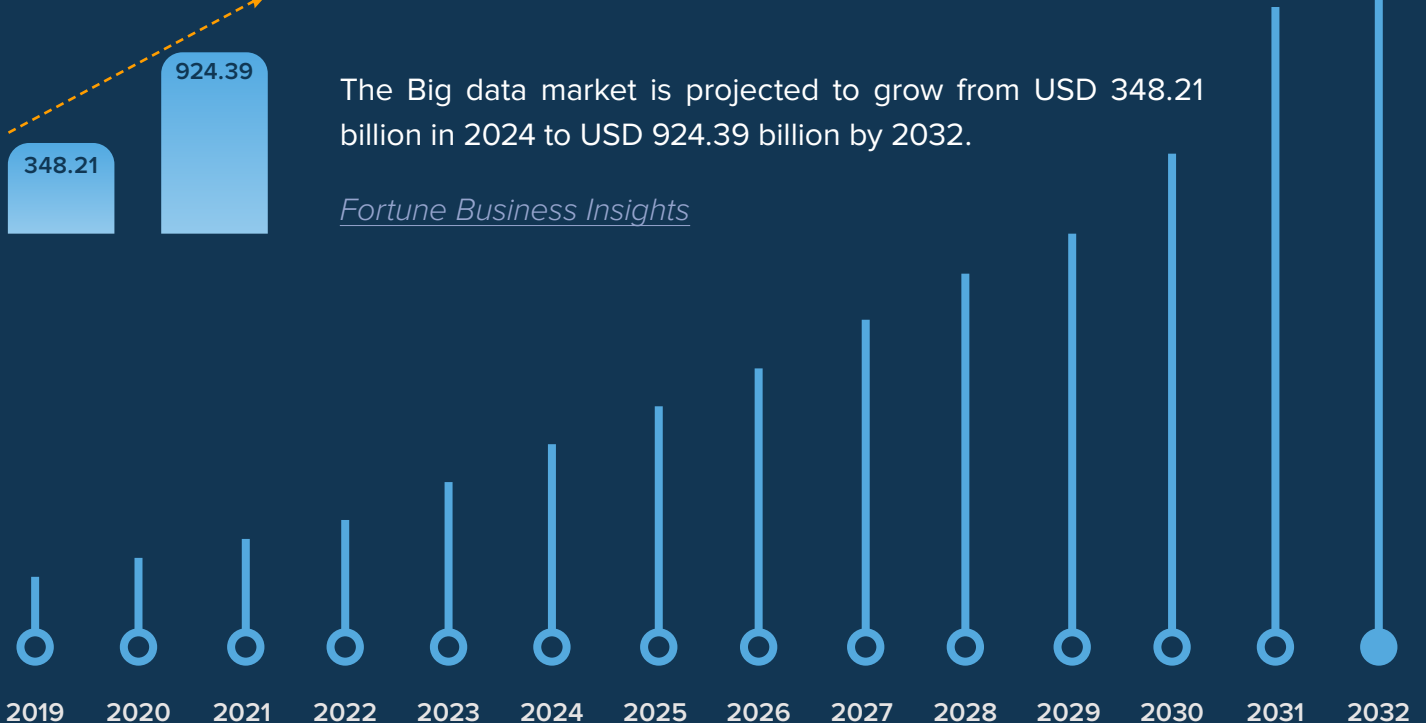
of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. This massive, diverse, and unstructured data, which is impossible to process via standard software and databases, is called Big data.

Big Data is getting bigger



The Big data market is projected to grow from USD 348.21 billion in 2024 to USD 924.39 billion by 2032.

Fortune Business Insights



90% of the data in the world today was created within the last 2 years. We will likely generate **480 exabytes** of data per day by 2025.

Raconteur

This rate will become greater with the growing popularity of IoT (Internet of Things) devices. Nowadays, data management is becoming a critical differentiator that separates market leaders from all others. Most enterprises face Big data, which is so large that it is impossible to process it using traditional software tools. Forward-thinking companies actively crunch their high-volume unstructured data to get a competitive advantage and find new business opportunities.

Today's high-end technologies make it possible to realize the value of Big data. For example, retailers, financial institutions and other B2C organizations can analyze the behavioral trends and social media activity of each customer and provide personalized product offerings; they can monitor customer satisfaction with company's products and services and take prompt marketing actions having sentiment analysis in place.

Data-powered predictive maintenance tools empower proactive business strategies that avert costly equipment downtimes and increase production capacities. According to Deloitte, this usage of Big data increases equipment uptime by up to 20% by predicting unexpected failures.

In a current data-laden world, BI reporting is another indispensable tool for modern businesses that helps companies make better decisions and take heed of all incoming insights. According to [Grand View Research](#), the global data analytics market size was valued at USD 49.03 billion in 2022 and is projected to grow at a compound annual growth rate (CAGR) of 26.7% from 2023 to 2030.

Oil and gas companies can take the output of sensors in their drilling equipment to make more efficient and safer drilling decisions.

Big data is a trend across business and IT, which refers to new technologies that can analyze high-volume, diverse data from traditional and digital sources inside and outside the company. Leveraging Big data analytics leads to more confident decision-making, which means greater operational efficiencies, cost, and risk reductions.

What is Big data?

Big data relates to data creation, storage, retrieval, and analysis, which are remarkable in terms of volume, velocity, and variety.

Three key differences between analytics and Big data:



Volume

so big, we need
new metrics



Variety

an abundance
of resources



Velocity

real-time
processing

Volume

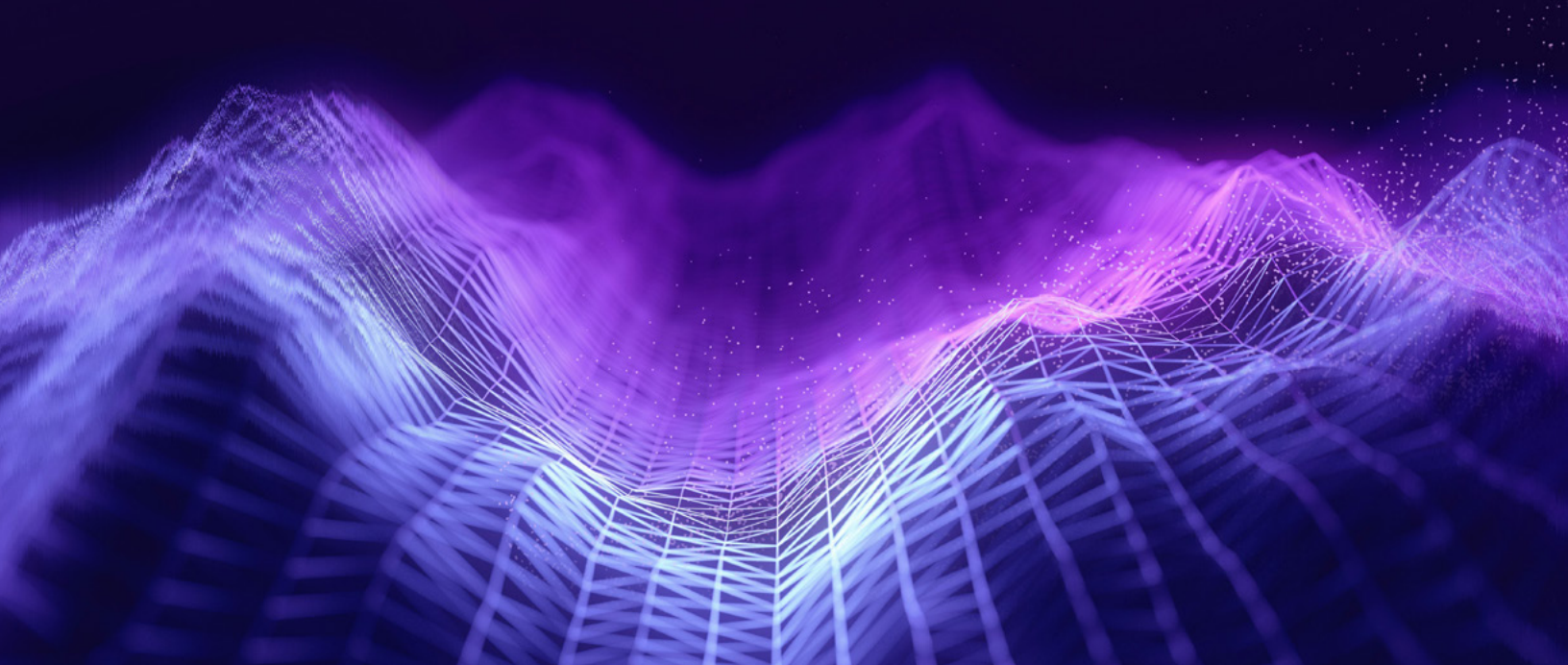
Massive volume of data is contributed by many sources of constantly updated data containing financial, environmental, location, and other information - transactions, social media, use of smartphones, and Internet of things. For example, Facebook produces 4 new petabytes of data every day; a Boeing 737 generates 240 terabytes of flight data during a single flight.

Variety

Data today comes in different formats: geospatial data, 3D data, audio and video, and unstructured text, including log files and social media. Managing, merging, and analyzing different varieties of data is a challenge for many organizations.

Velocity

Data is streaming in at exceptional speed and should be timely processed. Clickstreams and ad impressions capture user behavior at millions of events per second; high-frequency stock trading algorithms reflect market changes within microseconds; machine-to-machine processes exchange data between billions of devices.



What does this all mean?

It means that globally, companies are turning to Big data strategies to gain an edge over their competition. They realize that good business decisions are now data-driven and not intuitive. They analyze data to better understand and reach their customers, develop new revenue streams, and improve operational efficiencies.

Big data adoption grows at different rates in each vertical industry. Such markets as retail, financial services, telecommunications, and media are making considerable investments to effectively use their data to drive value.

The reason behind these verticals being the forerunners is that they have a lot of customers generating plenty of data, and a continuous need to keep customers happy so as not to lose them.

For example, the widespread use of increasingly granular customer data can enable retailers to improve the effectiveness of their marketing and merchandising. Data analytics applied to supply chains and operations will continue to reduce costs and create value and new competitive advantages for growing retailers' revenue.

BIG DATA AND GENERATIVE AI: How they align and complement each other

Generative AI and Big data have a significant relationship, as Big data plays a crucial role in the development and functioning of generative AI models. Here's how they're interconnected:



1. Training data:

Generative AI models, such as GPT-4, DALL-E, or any other, require extensive datasets to learn and generate new content. Big data provides the large and diverse datasets needed for training these models. The richness and variety of Big data enable generative AI to create more accurate, diverse, and contextually relevant outputs.



2. Improving model performance:

The quality and breadth of the data used to train generative AI models directly impact their performance. Big data helps improve the models' ability to generate high-quality text, images, or other content by exposing them to a wide range of examples and scenarios.



3. Training efficiency:

With Big data, generative AI models can be trained more effectively and efficiently. Larger datasets allow for better generalization, reducing the likelihood of overfitting and enabling the models to handle a broader array of inputs and generate more nuanced and sophisticated outputs.



4. Fine-tuning and personalization:

Big data can be used to fine-tune generative AI models for specific applications or user needs. For example, a generative AI model trained on a specific dataset can be further customized using domain-specific data to improve its performance in particular areas, such as legal documents or artistic styles.



5. Real-world applications:

In practical applications, generative AI models that utilize big data can create personalized content, such as tailored marketing messages or custom-designed products, based on patterns and preferences identified in large datasets.

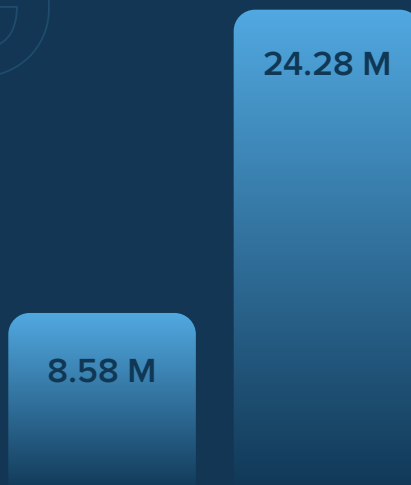
Big data and AI in banking and finance

According to Gartner, Big data in the financial industry has the highest level of opportunity due to the high volume and velocity of data available. Globally, Financial

Services and Banking are taking the lead in applying progressive Big data technologies and data science techniques, followed by Telecommunications and Retail.

"The Big data Analytics in Banking Market size is estimated at USD 8.58 million in 2024 and is expected to reach USD 24.28 million by 2029, growing at a CAGR of 23.11% during the forecast period (2024-2029)."

Mordor Intelligence



There are multiple internal data sources in banks and other financial institutions - including relational databases, XML data, Data warehouses, and enterprise applications such as ERP and CRM. Banks also have a large amount of external data about their customers

in the form of website visits, tweets, Facebook wall posts, searches, streams, videos, etc. This huge amount of data needs to be stored, processed and analyzed to help banks solve real business problems that banks face nowadays.

Typical banking sources of Big data include:



Customer bank visits



Call logs



Web interactions



Credit card histories



Social media



Transaction type



Banking volumes

Happily, technological advances and their reliability make it easier for banks to apply them to solve high-impact business problems. Data scientists now play a crucial role in applying Big data tools and mathematical algorithms to each specific business problem. Undoubtedly, banks and financial institutions of all sizes, shapes, and forms need to incorporate data science into their operating models.

What are bank business problems that data science can actually solve? They are multiple. Application of Big data in banking and finance varies from marketing (marketing campaigns efficiency, next best offer, personalized messages), through operational efficiency solutions (automated loan approvals, underwriting, and customer service) to fraud detection and risk management.

How generative AI and Big data are reshaping business problems in finance



Data-driven insights and decision-making

By utilizing Big data, Generative AI can develop predictive models, simulate financial situations, and produce tailored content such as financial advice or investment suggestions.



Risk management and fraud detection

AI can create scenarios to evaluate the strength of risk models and simulate potential fraud patterns.



Financial modeling and forecasting

Historical financial data, market trends, and economic indicators are analyzed to build models for predicting future trends, pricing, and market movements.



Trading opportunities

AI can use historical price data, trading volumes, and economic indicators to generate and test trading strategies by simulating different market conditions.



Regulatory compliance and reporting

Banks are required to manage and report large volumes of data to comply with regulations. Big Data tools help aggregate, process and analyze this data for compliance. GenAI can automate the generation of reports and regulatory filings. It can also help in scenario analysis to ensure compliance under various conditions.



Customer support and virtual assistants

AI-powered chatbots and virtual assistants can generate personalized responses, handle complex queries, and provide financial advice based on customer data (including chat logs, emails, and call transcripts).

"Banks need a modern technology architecture to transform with agility. It's important to have applications in the cloud, but when it comes to product development, it's also important that key workflows are built cloud-native. This enables automation and saves significant time."

EY

Also, analysis of social media helps banks predict customer churn. As the study by EY shows, **63% of customers** in the United States trust online personal networks and communities on choosing various banking products.

Moreover, 45% of customers comment on social media channels on the quality of service they received.

The ability to monitor customer sentiment gives banks early signals and allows them to be proactive in improving the customer experience and their engagement with the brand, thus saving costs and preventing revenue loss.

Analytics techniques can also play a significant role in fraud detection - allowing organizations to extract, analyze, and interpret business data to increase the probability of fraud and implement effective fraud detection systems.

All in all, Big data opens huge opportunities for banks and financial institutions. According to EY predictions, by 2030, banks will deepen

their personal connections with customers via data analysis techniques that might seem fantastic by today's standards.

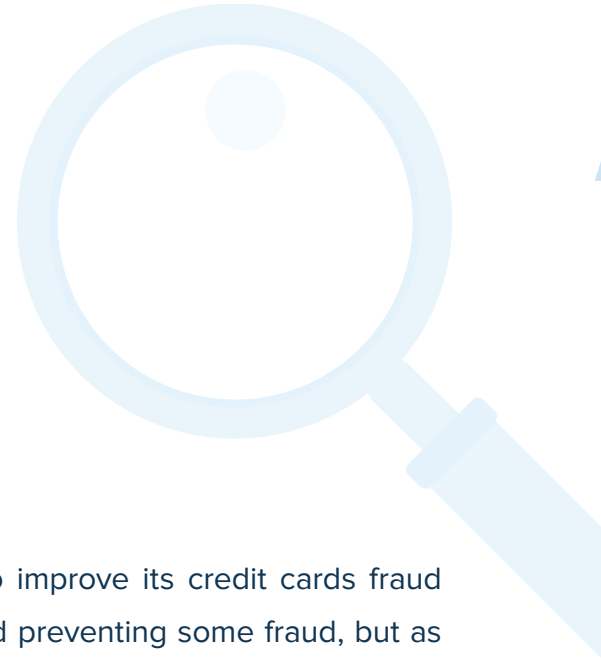
Although financial institutions have increased their awareness about data-field tools, some analytics applications are still in their nascent stage.

Therefore, in coming years, banking institutions will continue to look into the data-driven ways of customer engagement and interaction. Personalized services across channels will continue to enhance the consumer experience by demonstrating the interest of financial institutions for evolving customer needs.

BANKING AND FINANCE. CASE STUDIES

Case study

#1 Fraud detection



BUSINESS PROBLEM

One of the leading European financial institutions needed to improve its credit cards fraud detection capabilities. The existing process was detecting and preventing some fraud, but as the size and type of fraud varied and changed correspondingly, the bank required a new, more sophisticated system to provide efficient protection for the bank's customers.

SOLUTION PROVIDED

[InData Labs](#) offered to build an intelligent system that would detect fraud in real-time. The system applies the data analysis approach to create patterns in each client's behavior based on all of their previous live transactions. Then, it enriches this information with third-party data (e.g., social media data, geo-location) - such as geo data taken from the bank's mobile application. Transactions that do not fit into the cardholder's profile are marked as suspicious. Cardholders' profiles are updated with every single new transaction. The system uses self-learning techniques to constantly adapt to any changes in cardholders' behavior.

BENEFITS

- Fraudulent transactions are successfully distinguished from legitimate transactions
- Fraud prevented or minimized
- Decreased operating costs thanks to an automated approach.

Case study

#2

Contact center efficiency optimization

BUSINESS PROBLEM

A mid-size European bank faced the problem of rapidly rising call center operating costs and declining customer satisfaction. The bank's agents often had to make multiple transfers and switches of calls, with sometimes no result for the person who was calling. It became crucial for business to improve the contact center operations.

SOLUTION PROVIDED

To solve these problems, [InData Labs](#) developed a custom [AI-powered virtual assistant](#) for a bank institution that significantly improved operational efficiency and customer satisfaction.

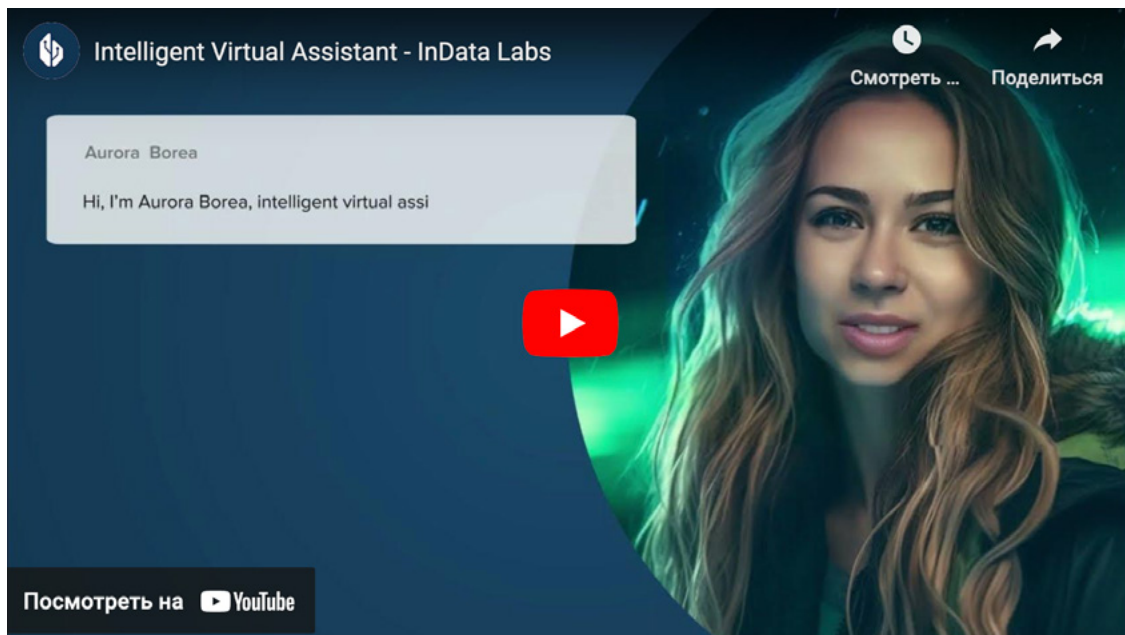
Aurora Borea is a proprietary intelligent chatbot of InData Labs designed to streamline query resolution, improve customer experience, and allow human agents to focus on more critical tasks with your advanced virtual assistant tool.

The solution stands out due to its sophisticated NLP and ChatGPT capabilities, enabling it to understand and respond to user queries with remarkable accuracy and context-awareness. This allows users to interact with the assistant as they would with a human, making it an invaluable tool for a range of applications, from managing schedules and automating repetitive tasks to providing in-depth data analysis and insights.

The virtual assistant's key feature is its ability to integrate seamlessly with various platforms and software, including CRM systems, project management tools, and communication apps. This ensures that users can streamline their workflows and access information without having to switch between different tools.

InData Labs has also ensured that Aurora Borea is designed with security and privacy in mind. The assistant incorporates robust encryption and data protection measures, giving users confidence that their sensitive information is secure.

Overall, Aurora Borea represents a significant advancement in virtual assistant technology, combining the latest in AI research with practical features that enhance productivity and ease of use. To learn more, please click to watch [this video](#):

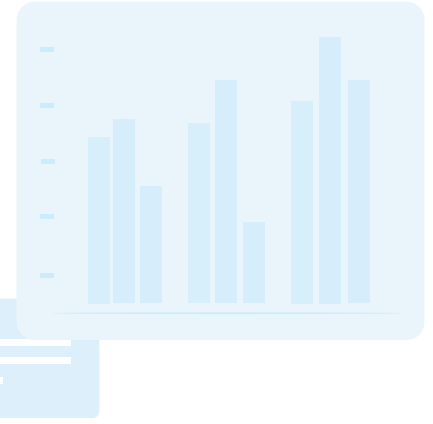


BENEFITS

- Instant response and 24/7 availability
- Multilingual support
- Automated contact center support
- Agility and scalability of business processes
- Reduced operating costs
- Improved cross-sales and up-sales success rates
- Increased customer loyalty.

Case study

#3 Customer churn analysis



BUSINESS PROBLEM

A large US financial institution confronted a serious issue - it was losing customers in favor of a competing bank. So, its management needed to identify the customers that were most likely to churn and offer them superb customer service in advance to keep them loyal.

SOLUTION PROVIDED

[InData Labs](#) developed a customer churn prevention system delivering analytics on each client in real time. In this decision-based system, mathematical algorithms and machine learning techniques are applied to historical data on lost customers. Based on this data, a churn model is developed.

The system identifies behavioral patterns and applies them to existing customers. Each customer is given a score that measures potential attrition. Moreover, the system models every situation and gives recommendations on the next best action for the customer to prevent churn.

For example, the bank's client used to get a salary from a bank account on a certain date of each month; however, these transactions stopped. The system analyzes it in real time, assuming that the client has changed or lost the job. To avoid the client's leaving the bank to use his/her new employer's preferred bank, the bank is able to offer him adjacent banking products. To verify if the client has lost their job, the bank gives them a courtesy call. As a result, the customer stays loyal to the bank and is ready to keep using the bank's products later on.

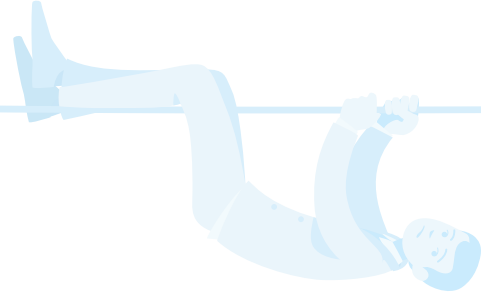
BENEFITS

- Upgraded customer service
- Increased customer loyalty
- Reduced attrition rates
- Identification of profitable customers to approach them directly.



Case study

#4 Risk management



BUSINESS PROBLEM

A large Western European bank wanted to implement a new system that would apply accurate methods to determine the credit risk of an individual or a legal entity. Basically, the bank needed a tool that would analyze credit applicants and determine their risk level with a very high probability in a very short period of time.

SOLUTION PROVIDED

To help the bank, [InData Labs](#) has developed a sophisticated analytical tool that is able to predict loan defaults with a very high probability. The high probability is reached thanks to a high-end predictive model developed by InData Labs` data scientists on the basis of enriched analysis of various data sources.

In terms of historical data, the system gathers and analyzes internal sources of information in the bank, such as credit reports and applications, repayment rates of credit applicants, and any information on default and recovery for borrowers. It also analyzes data from emails, website usage, and call centers. This sophisticated analysis is then enriched with data analytics from local credit bureaus and behavioral information from social media activities and other online sources of information (blogs, Google search, etc.).

As a result, the solution can provide the credit score for each applicant almost in real-time and with very high predictability of risk.

BENEFITS

- Higher credit scoring accuracy
- Improved credit decisions
- Credit risk-controlled and managed
- Healthy credit portfolio.

Case study

#5 Investment data management solution

BUSINESS PROBLEM

A group of US-based investment specialists struggled with managing investment data from various sources, such as Excel files and PDFs. They needed a solution to streamline data processing, enrichment, and integration to support their real estate investment platform.

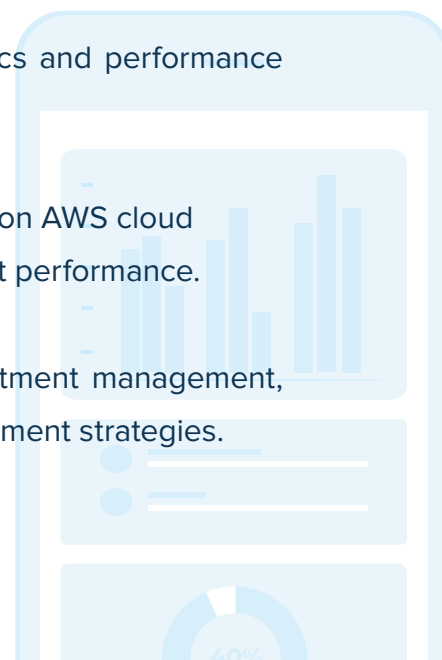
SOLUTION PROVIDED

[InData Labs](#) developed an end-to-end investment data management solution that extracts and processes users' investment data to simplify and add value to investment management activities. The team created data entry templates and built a data pipeline using AWS tools to streamline unstructured data processing and calculate key performance indicators (KPIs). The solution also featured data visualization through a Flask app and API integration with the client's web application. This allowed for secure, efficient data ingestion, helping fund managers make informed investment decisions with real-time metrics.

BENEFITS

- Streamlined processes: Simplified data management and processing
- Simplified investment management: Easy access to investment metrics and performance analysis
- Decreased data crunching time: Faster data processing and analysis
- Reduced error and security risks: Secure data storage and processing on AWS cloud
- Instant access to investment metrics: Real-time insights into investment performance.

The solution provided the client with a managed data service for investment management, meeting their business needs and enabling them to craft successful investment strategies.



SUMMARY

Big data and AI use cases in finance



ABOUT INDATA LABS

Our team delivers high-end engineering services & intelligent data analysis to achieve increased profitability of every business through constant insightful & data-driven management. Leveraging the latest big data technologies with a highly professional & talented team of data engineers, statisticians & mathematicians, we help our clients solve high-impact business problems in finance, customer acquisition, supply chain management, and risk analytics, to name just a few areas. Our core industry competencies are Finance, E-commerce, Supply chain & Logistics, and Digital Health.

Big Data Strategy Consulting

- Use case definition & prioritization
- Architecture design
- Road Map elaboration and Strategy report delivery

Engineering

- End-to-end deployment and management of Big Data platform
- Big Data integration services
- System engineering & technical support service

Data Science

- Customer analytics Predictive analytics solutions
- Generative AI & GPT integration services

More information about InData Labs services is available on the Web at www.indatalabs.com