



50 Ways

to Impact Your Business with AI



So you want to be an AI-driven enterprise?

Nearly every organization in the world right now is somewhere on the path to digital transformation. An important step in the journey is finding ways to leverage AI and Machine Learning to gain competitive advantage, enhance productivity and uncover hidden efficiencies.

Sometimes the most challenging part of getting started with these technologies is finding the right initial use case to work on. Not every use case is created equal. Some use cases are highly feasible, but not very useful. Others are high impact, but low feasibility. In order to find the best use cases for your organization, you must have right data, the right people involved in the project, and you need to be sure that everyone is clear on the desired potential outcome.

It's a common best practice to assemble a portfolio of potential use cases to explore. Lucky for you, RapidMiner offers a free [AI Assessment](#) and has hundreds of successful use cases to learn from. We've packaged up 50 of our highest-impact case studies that you can draw inspiration from. Consider this a veritable menu of enterprise AI options.

The following pages are called 'Impact Briefs' and they are real case studies of how AI helped drive progress in a RapidMiner customer environment. Some are anonymous to protect competitive advantages, and some are attributed with approval from our customers and partners. The one thing they have in common – all of these use cases have tremendous business impact. Many of them demonstrably increase revenue, most of them cut costs, and some reduce risk.

Some use cases are about a very specific industry but can still provide guidelines for businesses that aren't in the same industry. We've called this out on each use case under **'universal relevance'**

Which use case is the right one for your business to try and replicate? It's up to you. All you have to do is pick the one that looks most impactful and attainable, find some data, and get started.

What's holding you back?



By Industry

- 4 [Business Services](#)
- 8 [CPG & Retail](#)
- 17 [Discrete Manufacturing](#)
- 23 [Energy](#)
- 29 [Financial Services](#)
- 32 [Healthcare](#)
- 36 [Life Sciences](#)
- 42 [Process Manufacturing](#)
- 52 [Telecom](#)
- 55 [Travel, Transport & Logistics](#)
- 60 [Utilities](#)



Highlighted Impact Briefs

- 7 [Prioritize Marketing & Refine Brand through Deep Customer Insights](#)
- 9 [Forecast Supply Chain at Massive Scale](#)
- 11 [Improve Customer Experience with Expectation-Management](#)
- 19 [Improve Product, Marketing & Support Strategy with Unstructured Data](#)
- 30 [Automate Real-Time Reporting for High-Frequency Transactions](#)
- 40 [Generate Real-Time Public Sentiment Analysis](#)
- 48 [Optimize Process Control Variables to Increase Yield & Revenue](#)
- 55 [Understand Causes of Customer Churn](#)
- 58 [Improve Customer Service with Text Mining](#)
- 61 [Forecast Staffing with More Speed & Precision](#)



Business Services



Improve Resource Allocation for Better Revenue Conversion

Challenge

- Matches/places students with universities:
 - Help students avoid poor outcomes
 - Universities get high-caliber students
 - Pay 100% of recruitment and marketing costs
 - Marketing costs very high in some countries
- Poor (64%) conversion in key phase of process
 - From deposit placed to visa approved

Solution

- Predict the students who are at risk of visa rejection
 - High/Medium/Low Risk
 - Regular retraining of model keeps up with new cases
- Results surfaced in SFDC & Tableau
- Processes adjusted and optimized with prescriptive actions
- Enrollment services team can:
 - Prioritize high risk students in their workflow
 - Work with candidates to strengthen profile
 - Assist through visa preparation sessions

Impact

- Able to re-engineer process to resolve inefficient workflow
 - Deprioritized low and medium risk students
 - Save est. 135 hrs of work in month prior to visa issuance
- Enrollment team automatically provided assessment
 - Before every discovery call
- Every percent improvement in conversion
 - 13 students and \$300K in incremental tuition
- Improved conversion rate by 5.2%
- Helped drive incremental \$1.5M in January 2020

Problem type: Classification

Universal relevance: The use of AI can help re-prioritize allocation of resources to the most critical areas of need. When applied to long and expensive business processes, the impact can be enormous.



US STATE AUDITOR

Prevent Healthcare Fraud

Challenge

- Healthcare fraud can be challenging to detect
 - Even for intelligent medical providers & patients
 - Fraudsters well versed in remaining undetected
- Using a manual process to catch fraud wasn't working
 - Random samples selected for inspection
 - Less than 5% of all transactions were inspected
 - Time-consuming & limited resources

Solution

- Focused on specific behavioral patterns
- Easily prototyped & tested effectiveness of ML
- Supervised learning deployed:
 - Scan and flag high-risk fraud cases
 - In high volume, in real-time
- Integrated a variety of new sources:
 - Detection of fraud networks
 - Contextual clues that surround fraudulent acts

Impact

- \$20M Fraud identified
- New end-to-end detection & prevention process
- Able to identify and prioritize high-risk cases
 - Reduced time wasted with random inspections
 - Detect new fraud patterns
- New categories of fraud discovered

Problem type: Anomaly detection

Universal relevance: Fraud detection is usually manual and labor-intensive. At scale it can be a tricky problem to tackle. However, if you apply AI properly you can both save time and be more effective.



Prioritize Marketing & Refine Brand Through Deep Customer Insights

Challenge

- NFL team seeing declining home game attendance
 - Poor seasons & controversial player behavior
 - Overcrowded media market
- Needed to evolve brand and re-energize fanbase
- Partnered with 160over90 to:
 - Conduct market research
 - Evaluate brand awareness
 - Perform fan segmentation
 - Update key messaging

Solution

- Engineer features from survey data
- Develop a model of fan reactions to team messaging
- Create & validate a variety of segmentation models
 - Able to test dozens vs. classical approach of 3-4
 - Select the best fit model for the data
- 160over90 implemented "Ownable Asset Index" to
 - Analyze fan perceptions
 - Determine brand traits that align with fans
 - Isolate attributes that change perceptions

Impact

- Identified 4 distinct segments among NFL fans
 - Prioritize groups for marketing and PR
- Discovered growth opportunity among female fans
 - Able to significantly grow engagement with females
 - Enjoy football more in social settings
 - More likely to become fans later in life
- Uncovered opportunity to expand in greater metro area
- Discovered most impactful brand traits
 - E.g., intimidating, aggressive, blue collar, never satisfied
- Campaigns won a Clio Sports award for creative work
- Increased engagement & sales among target segments

160/90

Problem type: Classification

Universal relevance: Even if you think you know your customers, there's always new levels of insight you can generate by analyzing their preferences, behaviors, and perceptions. These insights will directly translate to more impactful campaigns and better branding.



CPG & Retail





Forecast Supply Chain at Massive Scale

Challenge

- Need to optimize supply chain forecasting
- Over 4k locations created scalability issues
 - Each location requires individual model
 - Resource and time intensive
- Shared data science resources are limited
 - Ecosystem of projects expanding rapidly
- Suffer from too many unique variables causing:
 - Food spoilage
 - Idle labor and overtime
 - Strained supplier relations

Solution

- Combined native operators & open source libraries
 - Leveraged RapidMiner's extensibility
- Able to leverage RapidMiner Server Scalability
 - Horizontally
 - Vertically
- Predictive models prescribe based on
 - Demand history
 - Previous and upcoming promotions
 - Holiday dates

Impact

- Prescribing accurate 8-week supply chain forecasts
- Optimize for:
 - Food waste
 - Labor pool usage
- Model runtime saw performance boost of 10x
 - Allows 4,000+ locations to receive forecasts

Problem type: Time series forecasting

Universal relevance: If your organization is building AI solution in code, like Python or R, then you likely suffer challenges managing deployments (runtimes, configurations, etc). Multimodal AI platforms can unify the development and deployment process across teams, while lowering costs and improving performance.





Create a Data-Driven Product Portfolio Strategy

Challenge

- Manages thousands of product SKUs
- Large volume makes it incredibly laborious to:
 - Analyze customer perception of products
 - Find factors that influence reviews & ratings
 - Isolate product review outliers

Solution

- Collect and aggregate data from many sources
 - Data repository for all company products
 - Includes competitor products
- Text analytics and topic recognition at scale
- Easy to find complaint-causing characteristics :
 - Quality
 - Packaging
 - Sensory characteristics
- Model fed to dashboard for lines of business

Impact

- Clearly identified high-performing products
 - Found undermarketed products
 - Optimized promotional strategies
- Faster fixes on packaging and quality issues
- Created a powerful competitive advantage:
 - Understand customers at scale
 - Agile adjustments to product strategy

Problem type: Text analytics & topic mining

Universal relevance: Even if your business doesn't manage thousands of SKUs, reading and understanding everything people are saying about your company or product is challenging at scale. Data science can help!



Improve Customer Experience with Expectation-Management

Challenge

- On-site wait time is too high during busy periods
 - Must communicate accurate times to customers
 - Poor projections impact customer experience
 - Customers abandon reservations due to frustration
- Heavy check-in staff overhead
- Self-service kiosk pilot unsuccessful
 - Too many variables for previous model to handle
 - Lack of model explainability made it hard to tune
- Brand and revenue impacted by poor experience

Solution

- Visual/collaborative platform for model building
- Automatic feature engineering found hidden insight
 - Surprising predictive variables for wait time
- Quickly prototyped and compared multiple models
- Model Ops used to see and prevent drift
 - Compare performance of champion/challenger
- Predictive models drive automation of process
 - Real time predictions delivered via web services
 - Integrated with kiosk interface

Impact

- Incremental revenue increased
 - Higher utilization of facility (food, beverage, etc.)
 - Fewer abandoned reservations
- Customer perception and loyalty boosted
 - Able to plan around more accurate wait times
- Realize the ROI of self-service kiosk investment
- Allocate resources more efficiently and cut staff overhead
 - Reducing hostesses by 2 employees/location/weekend
 - All US locations for 50 "weekends"/year
- Resulted in staff cost savings ~\$2M

Problem type: Regression

Universal relevance: Sometimes aspects of your product or service delivery are out of your control, which can lead to bad customer experiences. Accurately forecasting customer service times with AI mitigates risk and improves customer satisfaction.



ECOMMERCE RETAILER

Create Powerful Cross-Sell & Up-Sell

Challenge

- Acquisition cost of new customers is high
- Upsell of existing client base has potential to drive marketing ROI
- 1000s of catalog items create analysis challenge
 - Time consuming slow, resource intensive
- Hard to display targeted upsell ads in real-time

Solution

- AutoML preps data and compares 100s of models
- “Explainable” model output creates insight for anyone
 - Group customers with similar buying patterns
 - Identify why clients purchase additional products
 - Discover new patterns, trends, and insights
- Predictive models drive process automation
 - Prescribe the best course of action
 - Integrate with existing business processes
 - Deliver real time targeted advertising
 - Banner ads, emails, etc.

Impact

- Marketing drives better ROI
 - Can call out actual data points driving business
- Business makes more agile data driven decisions
- Can automate upsell activities
- Potential to drive ASP 2% or higher
 - Based on prior use case engagements

Problem type: Market basket analysis

Universal relevance: Shopping cart analysis and targeted advertising with AI has been the recipe for success for eCommerce industry titans. Leverage AI in your business to improve cross-sell and up-sell effectiveness.



Increase Sales with Insight From Reviews

Challenge

- Thousands of online reviews of dealerships
- Large volume makes it incredibly laborious to:
 - Analyze Dealership perception of products
 - Find factors that influence review ratings
 - Isolate product review outliers and upsell
- Analytics & subject matter experts do not code

Solution

- Collect and aggregate data from many sources
 - Data repository for all company products
 - Includes competitor sites with web scraping
- Surface complaint-causing issues related to:
 - Quality
 - Price
 - Experience
- Text analytics dashboard for topic recognition
- Report on dealerships with best/worst experience

Impact

- Faster fixes on negative dealership feedback
- Optimized promotional strategies
 - Identified highest performing products
 - Discovered undermarketed products
 - Upsell opportunities surfaced
- Created a powerful competitive advantage:
 - Understand dealerships at scale
 - Responsive adjustments to product strategy
- Increase sales to dealerships = \$500k

Problem type: Text analytics

Universal relevance: The impact of customer experience on revenue is hard to dispute, but it can be even harder to measure and analyze. With a mix of web scraping and topic recognition, you can monetize valuable insight from publicly-available information.



Optimize Service Distances to Maximize Delivery Margin

Challenge

- Costly partnerships with food delivery services
 - Take a 20% cut of the meal revenue
- Delivery area is crudely configured
 - E.g., 5-mile radius
 - Not based on most profitable regions
- Delivery options cannibalizing in-store pickup
 - Higher margins
- Must maximize radius w/o sacrifice freshness

Solution

- Predict optimal delivery area based on:
 - Demographics
 - Distances to restaurant
 - Previous orders
- Update the model based on factors that affect delivery:
 - Weather data
 - Traffic patterns
- Automatically update the model monthly:
 - Accommodate demographic changes
 - Reflect seasonality
- Report in MicroStrategy: delivery margin optimization

Impact

- Optimized delivery areas for each restaurant
- Improved same-store sales by 1%
 - Capturing more orders in optimized regions
- ~£1M to restaurant chain in first quarter
 - Across 1300 regional locations
- Management gains insight to delivery margins

Problem type: Prescriptive optimization

Universal relevance: In today's world of buy now/deliver now, customers expect their orders faster than ever, which drives up costs. Optimizing delivery with AI is essential to nearly every business in the world right now.



LARGE NORTH AMERICAN RESTAURANT CHAIN

Improve Customer Loyalty

Challenge

- Delivery chain looking to drive growth & expand
 - Tight margins
 - Intense competition
- Suffered from “one & done” customer segments
 - Often from online and mobile orders
 - Customers are fickle and rich with choices
 - Constant need to nurture and re-engage
- Must coordinate touch points to influence buyers
 - Physical, online, and mobile
 - Disconnected today

Solution

- Use vast amount of captured customer data
- Enrich the data
 - Geocode
 - Demographic
- Predictive models prescribe best offer based on
 - Store location
 - Customer profile & preferences
- Models deployed & managed with Server

Impact

- Posted best same-store sales growth ever
- Out-performed Wall Street expectations
- Over the same time period, top competitors:
 - Reported slowed revenues
 - Showed decrease in same store sales
- Generated more repeat orders
- Increased customer lifetime-value
 - 6% place another order within 30-days

Problem type: Text analytics

Universal relevance: The impact of customer experience on revenue is hard to dispute, but it can be even harder to measure and analyze. With a mix of web scraping and topic recognition, you can monetize valuable insight from publicly-available information.



Optimize Supply Chain SKU Management

Challenge

- # of SKUs has grown exponentially
 - 1500+ SKUs
- Too many variables for a human to manage
 - Extensive cost to hold/manage each SKU
- Downstream impact of SKU removal not always clear:
 - Volume doesn't tell the whole story
 - May be profitable items associated with a SKU
 - Profitability can be hard to gauge
 - Difficult to properly analyze these implications

Solution

- Text mining to group similar products
- Cluster the high and low sales revenue SKUs
- Cluster the high and low profit margin SKUs
- Auto correlation identifies low value SKUs
 - Replace them with high-value SKUs
- Model simulator shows impact of changes
 - SKU elimination
 - Using replacements for low value SKUs

Impact

- Holding cost per SKU \$40,000 a year
- Total holding costs of \$60 million
 - 1500 SKUs
- 5% SKU reduction delivers savings of \$3,000,000
 - Avoid downstream complications

Problem type: Text analytics & clustering

Universal relevance: Most organizations can benefit from some level of portfolio optimization. Volume doesn't tell the full story and profitability implications can be difficult to gauge without predictive analytics.



Discrete Manufacturing





Avoid Full Operations Shutdown with Predictive Maintenance

Challenge

- Must reduce plant out-of-service times
 - Results directly in lost revenue
 - Reduce unnecessary service crew travel costs
- Predict life-time of factory components & machines
- Predict machine failures that result in plant shutdown
 - Service needs before they become problems
 - Optimize maintenance schedule & crew utilization
- Anticipate needs for replacement components
 - On-hand as needed, without extra carrying costs

Solution

- Unify data in end-to-end tire lifecycle
 - Raw material to finished product
- Range of data sources in their models:
 - Sensor data from the plant operations
 - Log entries
 - Error and failure messages
 - Repair and maintenance service reports

Impact

- Drastically reduce risk of shutdown as result of:
 - Critical equipment failure
 - Parts for repair being unavailable
- Each avoidance \$20+ Million per/day cost
 - Likely to avoid 1-2 shutdowns per year

Problem type: Predictive maintenance

Universal relevance: Simple repairs and maintenance can have massive downstream implications. While 'disaster scenarios' may be rare, proactive avoidance with AI can keep them to a minimum or eliminate them entirely.



Improve Product, Marketing & Support Strategy with Unstructured Data

Challenge

- Reduce cost of post-sales support
 - With no sacrifice to customer experiences
 - Both with customer service & the product
- Customer support team struggled to:
 - Determine reason for calls
 - Identify a fit with existing solutions
- Strategy for betterment involves:
 - Improving the content on the website
 - Customers self-serve instead of calling for help
 - Better preparing agents for quick resolution of calls
 - Build products that require less overall support

Solution

- Turned to unstructured data in 26 languages to:
 - Go beyond basic statistical summaries
 - Determine causality
 - Find deeper insights that structured data cannot
- Analyzed:
 - Call logs and email inquiries
 - Online product reviews – competitors as well
 - Social media comments
- ML classification analysis:
 - Helped identify and flag trends across mediums

Impact

- Classification provided the “why,” not just the “what”
- Call center armed with better information
 - Predict/understand the reason for call upfront
 - Resolve issues faster
- Support team can refine the content on its web site
 - Full coverage for common issues
- Identify trends in customer perceptions for marketing
 - Adjusted communication strategies to resolve
- Insights delivered to product dev and manufacturing
 - Build products that won’t need support at all

Problem type: Text analysis & classification

Universal relevance: Unstructured data is hard to analyze, but it can be packed with rich insight. If you unlock the power of these insights, the benefits will permeate across the whole organization.



MAJOR AUTOMOTIVE MANUFACTURER

Master Variable Demand

Challenge

- Quickly-changing auto industry requires adaptivity
- Providing dealers with a sub-optimal vehicle mix:
 - Increases cost – managing un-sold vehicles
 - Lost revenue when supply not equal to demand
- Traditional approaches still have value
 - Need an approach to complement, not replace

Solution

- Recommendation engine created:
 - Provides the best mix of cars
 - Optimized for each dealership
- Advanced predictions enhance legacy forecasts
- Platform adopted by 100s of citizen data scientists
 - Empowers analysts with broad range of skillsets
 - Supports ever-expanding needs and use cases

Impact

- Hundreds of dealer orders optimized
- Sales forecasts are substantially more accurate
- \$10M+ benefit in year 1
- \$50M+ more expected in year 2
- RapidMiner enhances productivity of all teams
 - Data scientists free to focus on complex projects
 - Data analysts can work above their skill level

Problem type: Forecasting & prescriptive optimization

Universal relevance: Every supply chain requires precision forecasting. Inadequate quantities or improper mix means lost revenue. Forecasting with AI makes variability in demand much more manageable across the supply chain.



Predict Product Quality with Audio Mining

Challenge

- Sound experts predict engine lifetime & problems
 - Based on engine sounds at end of production
- Expert skillset is high cost and limited availability
- Test only applied to small sample of car engines
- Even expert human predictions have variation
- Desire to improve predictions and results by:
 - Automating audio data analysis
 - Applying test to more or all produced engines

Solution

- RapidMiner Time Series extension
- Provides audio data processing and transformation
 - e.g. Fast Fourier Transformation
- Machine learning predicts failures and quality issues
 - Based on transformed audio signals
 - Automated feature generation & selection
 - Automated optimization selects best ML algorithm and parameter settings
- RapidMiner Server used to automate predictions

Impact

- Significant increase in number of engines that can be tested
- Reduction of testing costs
- Reduced variation in predictions
- Increase in predictive accuracy
- Reduced late-stage engine repair/replacements costs

Problem type: Time series & audio processing

Universal relevance: Leveraging AI to mimic human cognition allows you to scale labor-intensive efforts for common tasks.



Reduce Time-to-Market & Costs on New Products

Challenge

- Market demands more & more truck models
 - Ever shorter cycles & smaller product series
- High cost for iterations between product designer and assembly planner
- Complex input data:
 - Complex 3D product designs (CAD)
 - 1000s of components and parts
 - Textual descriptions of each part
- Complex prediction problem:
 - Not classification or regression
 - Prediction of complete assembly plans

Solution

- Combination of machine learning approaches:
 - Text analytics, complexity reduction, clustering, classification, and regression
 - Predict assembly times and plans: 80% accuracy
- Automated assembly time prediction for each design:
 - Product designer gets immediate feedback
 - No iterations/consulting assembly planners
- Semi-automated assembly plan prediction:
 - Planner does not need to start plan from scratch
 - Can select most likely fitting plan and adjust

Impact

- Bring products to the market faster
 - Lower prices/larger profit margins
- Reduced time-to-market:
 - No wasteful iterations between groups
- Lower product design costs
- Lower assembly planning costs
- Able to meet demand for smaller product runs

Problem type: Text analytics, clustering, classification & regression

Universal relevance: Even iterative, creativity-driven business endeavors, like product design, can be optimized through intelligent application of AI.



Energy





Increase Yields without Sacrificing Quality

Challenge

- Refinement requires high level of mixture accuracy
- Improper ratios:
 - Damage expensive equipment
 - Deliver poor quality products
- Relied on intuition and experience to maximize yield
- Precision requirements increase costs substantially:
 - Increased use of raw goods
 - More waste – costly disposal

Solution

- Model the entire lifecycle of the machinery to:
 - Predict product quality based on configurations.
- Modeled many configurations to optimize yields
- Process running continuously to refine optimization
 - Making millisecond adjustments to equipment
 - Real-time (sub-30ms) model response time
- Easily explain model to domain experts

Impact

- Reduced reliance on institutional intuition
- Insight delivered through real-time predictions
 - No more reliance on daily manual readings
- Optimized machine settings and performance
- Produced impactful volume-based results
 - Yield increased 2%
 - Product waste decreased

Problem type: Prescriptive optimization

Universal relevance: Increasing output is not usually a hard problem to solve. However, doing it without increasing input or sacrificing quality targets is a tricky challenge that can be solved with AI.



Optimize Legacy Forecasting Models

Challenge

- Legacy forecasting platform is manual
 - Did not allow exploration
 - No advanced analytics
- Slow speed of development and deployment
 - New use cases not being solved
 - Existing models not continuously improved
- Each team is using a different platform
 - No collaboration or team growth
- Penalized for either over/under forecasting
 - Legacy tooling only provides limited algorithms
- The team has limited to no coding capabilities

Solution

- AutoML and code-optional approach
- Re-built existing models in RapidMiner
 - Leverage advanced algorithms (deep learning)
 - Increase consistency & accuracy
- Non-coders able to prototype new projects
- Deploy models to run automatically
 - No need to rebuild models every week

Impact

- A reduction of cost from legacy systems
- Increase in accuracy of models
 - ROI of this part of project still TBD
- Additional 10 hours per users for new projects
- Live tracking of all models - available to all
- ROI of \$200K+ with potential for much higher

Problem type: Forecasting

Universal relevance: Many companies employ basic statistical forecasting models for optimizing critical supply chain or other business fundamentals. The 'old way' is semi-manual, inflexible, and hard to maintain. AI does it better.



Eliminate Major Health, Safety & Environment Issues



Challenge

- HS&E issues with geysers/steam eruptions
 - Occurs during delayed cooker unit operations
 - Common during coke cutting step
 - Isolated hot coke pockets cause eruptions
 - Can spread across 2-3 mile plant radius
- Domain experts could not predict the eruptions.
 - 80+ plant operating parameters
 - Amount of data too high to analyze
 - Variables too interdependent

Solution

- Narrowed down to the most predictive variables
- Dashboard for real time monitoring of units:
 - Predictions built based on observed eruptions
 - Using volumes of historical data
 - Model accuracy of 99.1%

Impact

- Model allows operators to take corrective action
- Improved plant safety & environmental impact
- Steam eruptions have reduced over 80%
- Improved safety and overall HSE Score Card

Problem type: Regression

Universal relevance: Even if you don't have regular steam eruptions in your workplace, you likely have health, safety, and environmental regulations you need to adhere to. Predicting outcomes that lead to violations will help you avoid them.



Make Smarter, Risk-Based Investment Decisions

Challenge

- Gas distribution pipeline network is critical part of ops
 - Includes 2M+ miles pipeline in the US
- Every failure carries a 'nightmare cost' with potential to:
 - Be extremely costly to fix
 - Disrupt nearby residents & businesses
 - Create bad PR
 - Present a major safety hazard
- Unable to identify areas with greatest risk of failure

Solution

- Turbo Prep & Auto Model: instant insight from raw data
- Mature the insights through additional model building
- Quickly and easily build multiple models:
 - Choose from hundreds of modelling approaches
 - Compare results and predictive power
- Confusion matrices offer transparency into comparison
- Streamlined identification of sections in need of repair

Impact

- Reduced risk of:
 - Significant property damage
 - Loss of life
- Cost savings through more efficient repair operations:
 - Increased leak prevention
 - Decreased risk of catastrophic damage
- Optimized risk-based investment decision-making

Problem type: Predictive maintenance

Universal relevance: Predictive maintenance can not only be beneficial for improving efficiency, but also for avoiding events that affect public safety and cause brand and reputational harm.



MAJOR CRUDE OIL & NATURAL GAS PRODUCER

Optimize Use of Raw Materials

Challenge

- Synthetic crude oil is a key part of product line
- Hard to find right balance of how much to “upgrade” oil:
 - Can drastically affect process and total cost
- Precision requirements increase costs:
 - Increased use of raw goods
 - More waste – costly disposal
 - Inefficient transformation process

Solution

- Modeled the entire lifecycle of the oil to:
 - Predict product quality based on configurations
 - Prescriptive recommendations to process control variables
- Process running continuously to refine optimization
 - Optimize different levels of oil “upgrades”
 - Real-time model response
- Easily explain model to domain experts

Impact

- Increased profits by optimizing raw materials
- Production costs: \$21.05 per barrel
- Nearly 400k barrels per day produced
 - 24/7 production, 365 days
- Projected at \$3.2M increase in profits per year
 - Based on first 6 months

Problem type: Prescriptive optimization

Universal relevance: How much do you leave on the cutting room floor? AI can help ensure you're not leaving too much or too little. Either can be costly.



Financial Services





Automate Real-Time Reporting for High-Frequency Transactions

Challenge

- Reporting requires ETL and complex data prep
 - Leveraging many data sources
- Data analysts build non-IT-supported solutions
 - Run on analyst-supported prod servers
 - Custom tools, including custom code
- Analysts burn capacity to support infrastructure
 - Creates technical debt
- Stability of reporting is high-risk
 - Analyst turnover will cause big problems

Solution

- RapidMiner allows users of all skills to collaborate on a single platform:
 - Data professionals, business analysts, and data scientists
 - Driving collaboration and standardization
 - Lowering TCO
- Approved process automated on Server
 - High frequency reporting for brokerage
 - Process monitoring and governance
 - Results exposed through Tableau

Impact

- Low barrier to create new reports and models
- Standardized process:
 - Streamlines employee onboarding
 - Reduces downtime with turnover
 - Reduces technical debt
- Centralized, governed, and portable repository
 - For code and processes
 - Includes searchable metadata
 - Data analysts manage easy maintenance
- Management team: greater visibility into the data and the work

Problem type: Automation & data integration

Universal relevance: Up to 80% of a data scientist's time can be spent on data preparation. Custom-built data pipeline solutions are hard to maintain and creates organizational bottlenecks. Automating ETL and data prep builds a foundation for success for more advanced analytics projects.



ACCOUNTING FIRM

Automate Daily Manual Tasks

Challenge

- Difficult to identify missing, repeated, or bad data
 - Inefficient, manual and time consuming
 - Huge percentage of monthly team efforts
 - 16 Hours Per Project
- Information fatigue leads to errors
 - Errors create more direct and indirect cost

Solution

- Extracted and compared data sets in a couple clicks
- Developed a series of predictive models
 - Find repeat variables
- Project driven by a small, junior team
- Model was easy to put directly into production
 - Process re-used across projects
 - Process shared among users
- Automated process scales to larger data sets

Impact

- Manual Time: 16-hour project to 45 automated seconds
- Reporting is both accurate and timely
- Increased team efficiency – more time for:
 - Focusing on more strategic work
 - Delivering more custom projects
- Easily repeat and reuse process for new projects
 - Organizational Efficiency
 - Happier Staff
- Improved Quality & Reduced Errors

Problem type: Data integration & automation

Universal relevance: Every business is challenged to integrate data from disparate sources to enable business-critical analytics. These data integration tasks are time-consuming and labor-intensive. AI and automation streamline these tasks and shorten time-to-insight.



Healthcare



Improve Patient Satisfaction Scores with Text Mining

Challenge

- Low patient satisfaction scores
- Insight into low scores trapped in survey data
 - Hundreds of patients daily
 - Extensive manual review work
 - Burdensome to the patient experience team
- Hard to know what drives large-scale change

Solution

- Quickly analyzed survey data with topic recognition
- Identified most common root-causes of complaints and bad scores:
 - Doctor presence and face time
 - Wait times
 - Quality of care
- Process automated: analysis updated real-time
- Prescriptive recommendations delivered

Impact

- Worked with clinical staff to implement changes
 - Quick action & immediate results
- Patient satisfaction scores increased by 40%
 - First six months
 - Based only on first wave of changes
 - Continuous improvement expected from real-time updates
- Able to identify new issues, as they emerge

Problem type: Text analytics

Universal relevance: It's hard to manage what you don't measure. But what do you do when something is hard to measure? Topic mining can make unstructured data measurable, so you can identify actions that drive meaningful change.



Reduce False Positive Rate to Save Lives

Challenge

- Difficult to diagnose sepsis early
 - Timely diagnosis is critical for survival
- Large contributor to in-hospital mortality
 - 50% of all in-patient deaths
- Significantly impacts patient outcome ratings
 - Overdiagnosis = unnecessary meds/beds
- Diagnostic criteria has high false positive rate
- Criteria with better true positive rates are too slow

Solution

- Extracted variables from dense historical data
- Developed a series of predictive models
- Efficient training and testing of multiple models
- Certified selection of a model for real-time validation
- Project driven by small, under-skilled team
- Model was easy to put directly in hands of clinicians
 - Deployed at multiple hospitals
- Automated and enhanced quality of tests
 - 4-point manual test to 20 variable auto-assessment

Impact

- Increased patient outcome rates
- Diagnostic balance between accuracy and timeliness
- Easily prove value over previous diagnostic method
 - Simultaneously provide model transparency
- Reduced false positives = significantly reduced costs

Problem type: Classification & automation

Universal relevance: Few businesses deal with life and death consequences, but false positives in any business are just inaccurate assessments. When you automate and optimize assessments with AI, you can reduce manual effort and improve effectiveness.



GLOBAL HEALTHCARE PROVIDER

Prevent Patient Readmissions

Challenge

- Facing steep costs for readmissions
 - Services 4,000,000 members per year
 - 14 percent readmission rate
 - Average cost of patient readmitted is \$15,000
- Hard to identify why patients are returning
- Need a better way to analyze and regulate

Solution

- Rapidly prototyped ML model based on:
 - Attributes
 - Behavioral patterns
 - Claims data
 - E.g. emergency room visits that are not emergencies
- Built and deployed supervised learning model to:
 - Scan for at-risk readmittance patients
 - Accurately flag high-risk patients
- Targeted marketing and education for members
 - Getting the appropriate level of care
 - Importance of following discharge instructions

Impact

- 4M members x 14% rate = 560k readmissions
- 560k x \$15k = \$8.4 billion per year cost
- Achieving 0.1% improvement in rate:
 - \$8.4M in cost savings
 - Faster resolutions, better outcomes
 - Healthy, satisfied members

Problem type: Classification

Universal relevance: Most businesses relish in repeat business. For healthcare providers, readmissions are detrimental to public health and their bottom line. Identify attributes and behavioral patterns that may lead to repeat customers, for better or worse.



Life Sciences



Optimize Marketing Channels with Segmentation

CLARKSTON
CONSULTING

Challenge

- Sales & marketing efforts lacked engagement
 - Need to improve digital channel performance
- Needed analytics-based physician segmentation
- Overwhelming amount of existing client data:
 - Physician attributes
 - Prescribing patterns
 - Much of it is purchased & under-leveraged

Solution

- Developed a data dictionary from multiple sources
- Transformed, profiled, and tested every data set
 - Deep understanding of features' effect on engagement
- Group doctors with unsupervised ML
 - Analytically-driven feature selection
- Analysis on groups' engagement across sub-channel tactics
 - Digital and face-to-face
- Delivered 2 complimentary ML models
 - Predict values for key attributes, where missing
 - Predict cluster for future physicians in database
- Prescribed specific tactics and timing for groups of doctors

Impact

- Justified and enhanced value of existing purchased data
 - Better understanding of how to predict engagement
- Delivered novel insights; new ways to improve data capture
- Learned new, more effective approach to segmentation
 - Superior to traditional manual categorization
- Uncovered better engagement patterns
 - New groups of customers with engagement potential
 - Effective combos of channel, tactic, timing, & content

Problem type: Segmentation

Universal relevance: Most businesses can only make guesses about how likely prospective clients are to engage with their marketing channels. Accurate engagement predictions allow re-prioritization of marketing budgets to ensure efforts produce better outcomes.



Improve Research Effectiveness with Text Analytics

Challenge

- Over 10,000 researchers in one organization
 - Struggle to collaborate at scale
- ~1000 new research papers submitted/day
 - New discoveries
 - New work completed
- Requires a team of "indexers"
 - Read, tag, categorize, & organize all research
- Manual tagging slows down collaboration

Solution

- Researchers upload work via web portal
- Text analytics eliminates manual tagging
 - Topic analysis performed
 - Categories and meta-tags applied to documents
 - Auto-routes to right area of research library
- Model offers high degree of explainability
 - Process managed by non data scientists
 - Easy to tune and update

Impact

- Reduce millions in cost
 - Mostly through improved efficiency
 - Reduced manual work
- Dramatic increase in cross-team collaboration
 - Noticed broadly across global organization
- Additional cost savings identified
 - Near real-time updates
 - Better visibility into 'related' projects
 - Eliminated duplicate research efforts

Problem type: Text analytics

Universal relevance: AI is often applied to tasks that need to be done better, at scale. In this case, AI is helping a large group of people index and collaborate more efficiently over a massive body of text. Other examples patent portfolios, intellectual property libraries.



Precision Forecast to Cut Supply Chain Costs

Challenge

- 1000s of drug SKUs - complex inventory planning
- Inventory is expensive and demand fluctuates
- Using previous month's sales as main predictor
 - Significant statistical error (>20%)
 - Error could be present in either direction
- Require precise prediction of future sales to:
 - Reduce logistical costs
 - Streamline internal processes
 - Optimize the supply chain

Solution

- Domain experts can develop & validate models
 - Predictive models were cost-sensitive
 - Optimized for positive business outcomes
- Identified new, important sales influence factors
- Created 360° view of factors driving future sales

Impact

- Newfound precision in sales forecasting
 - Across dozens of product lines
- Optimized supply chain & investments
- Decreased product waste
- Reduced unnecessary carrying costs
 - Contract right amount of storage space

Problem type: Regression

Universal relevance: You don't need to have 1000s of SKUs with expensive shelf-stability implications in order to benefit from more accurate and precise forecasting methodologies. AI can help cut your supply chain costs.



PHARMACEUTICAL

Predict Market Share Threats

CLARKSTON
CONSULTING

Challenge

- Facing threat of competitor drug
 - Marketed as better alternative to a client offering
 - Advantages from improper dosing of client drug
- Measuring adoption of competitive drugs is difficult
 - Particularly in existing client base
- Possess external data about the doctors' preferences
- Goal: target doctors likely to transition to competitor
 - Take proactive marketing action
 - Coordinated sales effort

Solution

- Leverage internal and external data sources for analysis
 - See each attribute's effect on new drug adoption
- Create likelihood assumption of the doctor's adoption
 - Gradient boosted tree algorithm selected
 - Internal client data to predict & validate assumptions
- Physician groups classified by "likelihood to adopt"
 - Low, medium, or high
- Further segmented medium and high segments:
 - High volume of patients
 - Those who are sub-optimally dosing their patients

Impact

- Action plan created to defend market share
 - Target specific groups of doctors
- Create lists for targeted sales effort
- Arm reps with critical information, including:
 - Further education on disease that drug treats
 - How to properly dose
- Channel optimization shows best way to reach audiences
- Improved organizational competency
 - Exploratory analytics approach to problem-solving
- Replicated model for other competitive markets

Problem type: Clustering & classification

Universal relevance: Any marketer knows the value of segmenting their audiences. Using predictive analytics you can segment based on 'likelihood' to take a specific action, which enables you to identify threats and opportunities in your market.



Generate Real-Time Public Sentiment Analysis

Challenge

- Developing and launching a new type of injection needle
- To mitigate risks of launching a costly item, needed data on:
 - Public's opinion on the company' and its products
 - Public's opinion on an existing comparable needle
- The data they were already collecting had some problems:
 - Labelling of raw data
 - Biased towards the vendor – no competitor data
- They needed a way to scrape, collect, analyze, & display:
 - Social media sentiment
 - Other online data (especially from forums)

Solution

- Execute and schedule web mining tasks
- RapidMiner web scraping crawls relevant webpages
 - Captures overall sentiment, common complaints, etc.
 - HTML data from online forums, Twitter, Facebook
 - Queried industry-specific databases
- Integrated with commercial web scraper software
 - Expands to Java scripts, pages with log-in credentials
- Integrated with database specialized in document handling
- Used ML to score the data
 - Intelligently find & cluster keyword combinations

Impact

- Web app displays results in a real-time, intuitive dashboard
- Dashboard includes a live search box:
 - Employees can type a specific word - product names
 - Dashboard instantly shows all relevant internet data
 - ML ensures only the most relevant info displayed
 - Incredible insights about public sentiment at fingertips
- Increased customer satisfaction
- Rich competitive intel on new products and their adoption
- Able to improve existing products based on sentiment
 - Avoid launching faulty products

Problem type: Sentiment analysis & text analytics

Universal relevance: Public opinion is commonly represented anecdotally, yet it dramatically alters business strategy. With the power of AI, it doesn't need to be. One employee described this solution as "CTRL+F? for public sentiment from across the globe." Who couldn't use that?



Process Manufacturing



Reduce Real-Time EHS Risk with AutoML

Challenge

- Shortage of classically trained data scientists
 - Process Engineers need self-service modelling
 - Constant flow of new and emerging use cases
 - Lack the experience or tooling to do it alone
- Must avoid shutting down plants due to EHS risk
 - Causes major reductions in output
 - Need optimum targets for process parameters

Solution

- Using AutoML Process Engineers:
 - Identified the variables with predictive power
 - Prescribed actions to reach optimum targets
 - Learned how to avoid environmental release
- Engineers empowered to perform tasks quickly
 - Turbo Prep for fast & guided data preparation
 - Auto Model for rapid, automated prototyping

Impact

- Identified a variable with 73.8% correlation
- Operators able to keep plant from going offline
 - Avoid subsequent administrative work
- Plant navigated an upset condition
 - Reduced EHS risk by 90%
- \$150k in gains for the incident
 - More than 6 incidents potentially avoided/year

Problem type: Prescriptive optimization & data science democratization

Universal relevance: The simple act of building a model can teach you invaluable lessons about your processes. If you put the power of rapid prototyping into the hands of subject matter experts, you'll empower them to use their knowledge to improve your operations in ways you can't imagine.



Optimize Yield by Avoiding Degenerated Equipment Performance



Challenge

- Equipment fouling disrupts Ethylene production:
 - Severely impacts plant productivity
 - Compounds energy costs
 - Increases maintenance costs
- Impossible to identify fouling indicators early
- Building correlations for predictions is difficult
 - Parameters are highly interdependent
 - Traditional statistical methods did not work

Solution

- Quickly pinpointed the most predictive variables
- Built dashboard for real time monitoring
- Helps predict fouling rate:
 - Normal Operation
 - Suspected Fouling
 - High Fouling/"Needs Immediate Action"
 - Immediately receive a predicted arrival time
- Completely automated scheduled maintenance

Impact

- Early identification of fouling decreased costs:
 - Dramatically reduced energy consumption
- More precise scheduled maintenance:
 - Decreased unplanned downtime
- Annual cost benefit of \$550,000+
- Increased productivity
 - Reduced downtime by 15%

Problem type: Classification & automation

Universal relevance: Nearly every business has some element of performance degradation over time. Understanding root causes of degradation and optimizing maintenance schedules can help maximize overall equipment effectiveness.



GLOBAL PAPER MANUFACTURER

Empower Citizen Data Scientists

Challenge

- Organization is embracing digital transformation
- Data silos across the globe, impeding progress
 - No way of sharing findings from plant to plant
- Different plants & regions using different tool sets
- 3 data scientists for entire organization
 - Only one per region, could not scale efforts
 - Lack of model explainability
 - Difficult for engineers without coding skills
 - Hard to tune and understand

Solution

- RapidMiner collaborative platform allows:
 - Engineers to jumpstart data science projects
 - Quickly create and compare multiple models
 - Continuously train and update models
- Share insights and deploy models globally
 - Eliminate data silos
 - Integrated into digital twin initiative

Impact

- Plant engineers empowered
 - Can address new challenges with ML
- Increase number of problems that can be solved
- Continuous improvement accelerates innovation
 - Drives global collaboration
 - Furthers digital transformation initiatives
 - Enhances investment in digital twin
- Time to resolve data science issues drop by 90%

Problem type: Data science democratization

Universal relevance: Data science democratization isn't just a tooling issue—it's an organizational issue. Providing subject matter experts with the ability to explore machine learning on their own, has a transformational impact on how cross-functional teams solve emerging problems.



Gain a Competitive Edge with Yield & Quality Optimization

Challenge

- Key product line represents ~\$100M revenue/yr
- Market dominance relies on high quality product
- Currently discard 25% due to high standards
- Digital cameras used for inspection
- Process engineers still must manually inspect
 - Labor intensive
 - Error prone

Solution

- Process data captured
 - Per second temp. in each zone of process
 - Combined w/ manual reports from engineers
- Model production lifecycle with digital twin
 - Predict # of deformities
- Prescriptive optimizer minimizes defects
 - Optimizes temp. for specific properties
 - Enhance quality progressively with ML
- Engineers trained in AI/ML for diagnostics

Impact

- Every correct prediction saves product
 - Reduces waste
 - Improves yield and reduces costs
- Up to 50% of discards avoided
- \$8M-\$12M conservative savings estimate
 - ~\$1M Month
- Maintain extremely high quality standards

Problem type: Prescriptive optimization

Universal relevance: AI augmentation of institutional knowledge can improve both output AND quality. This kind of impact creates sustainable market dominance.



MAJOR PETROCHEMICAL MANUFACTURER

Optimize Energy Costs

Challenge

- Ethylene plant relies on Crack Gas Compressors
- CGCs facing massive low efficiency issue
- Typical 500 KTA capacity ethylene plant:
 - CGC power accounts for 30 MWh of electricity
 - Equivalent to \$20 million per year
- Traditional management approach is ineffective
 - 100+ interdependent operating parameters
 - Relies on observation & domain expertise

Solution

- Identified 40 variables that drive plant efficiency
- Deployed models that can adjust parameters
 - In real time
 - Based on sensor data
- Dashboard for plant operators
 - Monitor model performance and predictions

Impact

- 5% reduction in power consumption
 - Equivalent to 1MW
- Turn around maintenance schedule extended
 - 6-8 month impact
- Total savings of \$1 million per year

Problem type: Prescriptive optimization

Universal relevance: Not every business has a single asset that consumes \$20M/year in energy costs, but they can add up no matter the industry. AI offers a fantastic way to be more green and save more money.





Optimize Process Control Variables to Increase Yield & Revenue

Challenge

- Competing with emerging, modern, global competitors
- Universal initiative to optimize 'recovery'
 - Recovery = yield produced from chemical reactions
 - Critical measure of process efficiency
- Must better understand:
 - What process variables influence recovery
 - How changes to variables impact recovery
 - Optimum targets for key process variables
- Operators can confidently make changes on the floor
- Plant Managers drive reproducible, long-term improvement

Solution

- Leveraged a wealth of unlabeled data
- Created a model to identify variables and trends
 - Help detect recovery imbalances
- Dashboard with predictions based on the model
- Delivered via reliable & intuitive mobile UI
 - Minimal training requirements
 - Approachable for any user
 - No barriers to consuming the insights
 - Simple to make changes on-the-fly

Impact

- Maximize output from limited ore reserves
- Annual production of 8M tons
- Previous YTD Recovery = 87.8%
 - Variations between crews as high as 3-5%
- Target to improve final product yield by 1%
 - Decrease variance between crews
 - No sacrifice to OEE or quality
 - ~56k tons of additional concentrate
 - **Worth approximately \$2M to the business**

Problem type: Prescriptive optimization

Universal relevance: The most important decision is the next one made by the boots on the ground. People-centric applications of AI empower workers to make better decisions, leveraging more information than ever to maximize business impact in any set of conditions.



THE WORLD'S LARGEST STEEL MANUFACTURING COMPANY

Reduce Re-work to Slash Costs

Challenge

- Steelmaking is complex and energy intensive
- Identifying defects early is critical
 - Spares energy expenditure of final steps
 - Product will be defective regardless
- Must find defects early & optimize production
 - Keep the cost of final output low
 - Stay competitive

Solution

- Leveraged a wealth of sensor data in aggregate
- Predictive models monitor casting & rolling processes
 - Detect anomalies early
 - Identify variables that cause defects
- Advanced time series feature extraction
 - Represent the shape of time series data
 - Enables detection of similar shape patterns

Impact

- Able to identify possible production problems
 - As early as possible
 - Explore them to understand them
- Custom dashboards help in a variety of ways:
 - Monitor results of defect detection process
 - Enables team to continuously improve model
- Less metal re-work



Problem type: Anomaly detection

Universal relevance: Not every business has 're-work' as part of their operations, but re-work is just a specific kind of bad outcome. Predict bad outcomes before they happen, so you can stop them at the source.



Prevent Exceeded Emissions & Critical Situations

Challenge

- Complex continuous production processes
- Detect anomalies as quickly as possible & alert
- Large variety and volume of data
 - 1000s of sensors, high measurement frequency
 - Factory & machine configuration data
 - Operator logbooks and other textual data
- Must learn from past issues and resolutions
 - Context
 - Success in quick issue resolution or prevention

Solution

- Access & ingest a wealth of sensor data
- Advanced time series feature extraction
- Time series analysis & forecasting
- Machine learning (classification) used for:
 - Automated detection of known issues
 - Automated detection of new issues with anomaly detection methods
 - Automated identification of similar past situations – decision support for operators

Impact

- Identify potential production problems
 - As early as possible – allows for exploration
- Faster response to critical situations
- Prevention of critical situations
- Reduction of risk & costs:
 - Fewer undesired emissions
 - Fewer equipment damages & failures
 - Lower repair costs
 - Fewer outages & related costs

Problem type: Time series & classification

Universal relevance: With the prevalence of IOT technology, many businesses have rich sensor data they're not fully utilizing. The analysis of this data can have massive, wide-ranging implications and should always be explored for AI/ML projects.



Eliminate Waste with Better Quality Control

Challenge

- Glass making is complex and energy intensive
- Identifying defects early is critical
 - Spares energy expenditure of final steps
 - Product will be defective regardless
 - 1-hour gap from production to inspection
- Must optimize production & find defects early
 - Keep the cost of final output low
 - Stay competitive

Solution

- Leveraged a wealth of sensor data in aggregate
- Predictive models monitor the casting process
 - Detect anomalies early
 - Identify variables that cause defects
- Advanced time series feature extraction
 - Explore shape of time series data distribution
 - Detect distributions that produce defects

Impact

- Identify & explore production problems early
- Custom dashboards help in a variety of ways:
 - Monitor results of defect detection
 - Enables team to continuously train model
- Prescriptive optimization adjusts furnace control
 - Based on sensor and model inputs
- Overall plant efficiency increased by 1%
 - Multi-million dollar annual per-plant impact

Problem type: Anomaly detection

Universal relevance: Not every business has a product as fragile as glass. However, every organization has processes that will occasionally produce bad outcomes or defects. Modeling can help uncover the root cause of bad outcomes so they can be remediated.



Telecom





VERIZON WIRELESS

Understand Causes of Customer Churn

Challenge

- Pre-paid churn rates higher than post-paid
- Predict churn-likely customers before next bill
 - Target them with remediation or marketing
- 32 million prepaid transactions daily
 - 40 different channels
 - 100+ transaction types

Solution

- Capture all real-time data and aggregate
 - Complex pre-processing process
- In-depth feature engineering process:
 - 200 attributes collected; 80 generated
- Compared 7 models – GBT selected:
 - Accuracy
 - Model confidence

Impact

- Gained actionable insights from data
- Identified leading indicators of churn for:
 - Pre-paid customers
 - Post-paid customers
- Were able to proactively identify 30% of churn
- Able to challenge fundamental assumptions
 - What was causing the churn?
 - How addressable is the problem?

verizon[✓]

Problem type: Classification

Universal relevance: Customer retention is a top priority for most businesses as the cost of acquiring new customers is high and constant churn creates a drag on profitability. Understanding why customers churn with AI delivers a clearer path to reliable results.



Eliminate Fraud to Increase Customer Satisfaction

Challenge

- Influx of customer complaints about SMS fraud
 - Messages can mimic a reputable source
 - May contain a virus or malware
- Case managers having difficulty acting at scale:
 - Identifying fraud sources quickly enough
 - Taking remedial measures to block them

Solution

- Data & metadata for all messages analyzed:
 - Sender
 - Message length
 - Frequency
 - Volume & more
- Leveraged Hadoop for enormous data stream
 - Capture, manage, & analyze in real-time
- Predictive models find spammers automatically
- Automation used to block once identified

Impact

- Learned new variables indicative of spammers
- Reduced nuisances for customers
- Results within a couple of weeks:
 - Massive reduction in instances of fraud
 - Slashed time to intervention
- Customer satisfaction metrics rose by 20%
- Complaint levels fell 50%

Problem type: Classification

Universal relevance: Customer retention is a top priority for most businesses as the cost of acquiring new customers is high and constant churn creates a drag on profitability. Understanding why customers churn with AI delivers a clearer path to reliable results.



Travel, Transport & Logistics



Improve Profitability by Predicting Delays

Challenge

- Late flight arrivals severely impact operations:
 - Disrupts catering services
 - Delays flow of flight crews between aircrafts
 - Jeopardizes gate availability
 - Delays connecting flights for passengers
- Identifying delays early makes it easy to adjust
- Delays are hard to predict & slow to identify

Solution

- Quickly pinpointed the most predictive variables
- Built innovative yet simple widget for flight ops:
 - Input specific flight numbers
 - Immediately receive a predicted arrival time
- Automated flight scheduling
 - Includes the predicted arrival times

Impact

- Discovered prime factors that create delays:
 - Which runways are utilized
 - How many passengers have connections
- Model accuracy optimized at time of take-off
- Annual cost benefits in the 10s of millions



Lufthansa
Industry Solutions

Problem type: Regression

Universal relevance: Every business experiences some kind of delays, which can cost you time, money and effort. Identifying delays before they happen gives you time to prepare.



Reduce Costs Caused by High-Value Asset Downtime

Challenge

- Looking to predict fleet maintenance needs
- Out-of-shipyards breakdown implies:
 - Grounding (loss of revenue)
 - Towing costs
- Time in shipyard carries high opportunity cost
 - Need to maximize time spent in shipyard
 - Execute repairs quickly, proactively
 - Avoid waiting for parts
- Spare parts storage is costly and inefficient

Solution

- Ingest data from diverse systems, including:
 - Error logs & messages
 - Onboard sensor data
 - Route schedule
 - Weather history
 - Maintenance reports
- Multiple models created to predict & manage:
 - Ship maintenance needs
 - Corresponding parts requirements

Impact

- Increased operational earnings 3% by:
 - Reducing unplanned maintenance
 - Increasing available chartered days
 - Optimizing spare parts storage worldwide
- Proactive maintenance program
- Full overhaul of communication strategy

Problem type: Predictive maintenance

Universal relevance: Most businesses have a critical asset on which their core product or service depends. AI can help you identify the best way to keep these assets in service and performing, so that your business runs unimpeded.



Improve Customer Service with Text Mining

Challenge

- Customer service criticized over responsiveness
- Overwhelmed team of two employees
 - ~ 500 messages per day, on average
- Messages require routing to other departments
 - Slow routing = even slower response time

Solution

- RapidMiner's text mining parses every message
 - Breaks down to key words and phrases
 - Helps prioritize messages appropriately
 - Improves case management
- Predictive classification drives automation
 - Auto-route messages to the right department

Impact

- The customer service team saves time
 - Reduced manual triage of incoming messages
 - No time spent routing cases to other depts.
- More time & resources to focus on 'hot' issues
- Negative social media sentiment down by 50%
 - From ~90% to ~40%



Problem type: Classification & automation

Universal relevance: Brand is everything. Streamlining customer service workflows creates a responsive, interactive brand for your business. Strong brands encourage future buying.



GLOBAL MINING FIRM

Reduce Fuel Costs

Challenge

- Each site hauls 200,000+ tons of material/day
 - Consistent and predictable hauling routes
- Experiencing wild variations in fuel usage
- Fuel accounts for up to 30% of energy usage
- Tried many things to optimize fuel
 - Paths, speeds, & investing in newer trucks
- Early modeling suffered from
 - Biasing
 - Overtraining

Solution

- Transform unstructured time series data
- Build model to estimate fuel burn rate per truck
 - Benchmark actual values against observed
- Model generalized well across the fleet
 - Doesn't overfit for certain trucks/truck models
- Push real-time data to user-friendly dashboard
 - Adjust parameters & explore results

Impact

- Found most important factors for consistent fuel burn:
 - Oil level & quality indicators
 - Based on sensor data and last tune-up
 - Load weight
- Able to precisely forecast fuel consumption
- Able to prevent trucks from performing poorly
 - Detect anomalies
 - Perform maintenance proactively
- Decreased operating expenses by millions per year

Problem type: Forecasting & predictive maintenance

Universal relevance: Unexpected variability makes it difficult to manage business performance. Understanding the causes of the variability leads to processes that improve consistency and generate reliable outcomes.



Utilities



Forecast Staffing with More Speed & Precision

CLARKSTON
CONSULTING

Challenge

- Call volume dictates staffing levels
- Existing volume forecasting was falling short
- Inaccurate staffing proved to be very expensive
 - Overstaff: Direct cost
 - Understaff: Indirect cost, bad customer service
- Looking to improve advanced analytics function overall to gain a competitive edge
- Needed project management guidance and assistance applying best practices

Solution

- Two forecast models using RapidMiner:
 - Short-range model using machine learning
 - Long-range model using statistical methods
- Integrated the forecasts with Qlik for reporting
 - Easy distribution for weekly ops meetings
- Regular knowledge transfer sessions to ramp up data literacy across teams
- Identified organizational gaps to address

Impact

- \$665K projected annual savings per call center
- Achieved average accuracy of 93% for 90-day forecast
- Automation frees up resources for other projects
- Formalized a repeatable process for:
 - Moving ML models to production
 - Maintaining models
- Identified key steps to making ML projects successful
 - Templates and best practices to be applied to other use cases

Problem type: Forecast & regression

Universal relevance: Labor is the largest variable cost for many industries. Forecasting staffing is both critical and hard to do because there are so many factors that can impact staffing demands. Machine learning helps drive more precision.

FirstEnergy™



Streamline Customer Service with Text Mining

Challenge

- QA team unable to:
 - Filter calls effectively
 - Flag for further action
- Call centers produce lots of rich data
 - Much of it is unstructured
 - Hard to analyze
- High volume of unresolved calls and complaints
- Faced tremendous customer satisfaction issues

Solution

- Acquire unstructured data from 3rd party transcription
 - Use as foundation of predictive models
- Unsupervised learning applied to unstructured text
- Call data labeled & reported based on:
 - Region
 - Urgency
 - Topic
 - Volume
- Deployment & automation of models in production
 - Managed with RapidMiner Server
 - Filters and routes calls for human review

Impact

- Reduced customer service complaints by 8%
- Reduced proactive unhappy customer callbacks
 - Labor intensive problem
- Identified problematic call centers and reps
 - Deployed proactive, targeted re-training

Problem type: Text analytics & automation

Universal relevance: Every organization has a 1-800-WE'RE SORRY number. If you manage this customer service hotline proactively with AI, it becomes a competitive advantage instead of a burden.



MAJOR UTILITY COMPANY

Execute a Revenue Assurance Plan

Challenge

- Too many late & defaulted customer payments
- Created major risk, revenue, and cost issues
- Forced to hold large amounts of cash on-hand
 - Unable be used for capital improvement
- Must sell debt to collection agencies
 - Usually for a very big loss

Solution

- Classification of customer base
 - Propensity to pay vs. ability to pay
 - Identify high risk to be late or default
- Automated workflows fed by the predictions
- Deploys best intervention strategy per customer

Impact

- Increased on-time payments by 4% overall
 - Mitigated financial risk caused by debt
- Better overall cashflow
 - More to invest in capital improvements
- Able to help under-privileged customers:
 - Take proactive intervention approach
 - Match with assistance agencies
 - Offer payment plans

Problem type: Classification

Universal relevance: Cash on hand is a critical part of any CFO's strategy. Gaining more insight into high-risk accounts receivable helps you develop smart intervention strategies, but it also allows you to make necessary adjustments to your revenue assurance plan.



Build a 'Monetized Risk' Approach to Maintenance

Challenge

- Electric grid is critical part of operations
 - Obligation to deliver reliable electricity
 - Maintain a massive grid infrastructure
- Major need to optimize repair investments:
 - Thin margins
- High risk of not fixing critical damage and wear & tear
- Natural disasters can wreak havoc
- Must know where is highest likelihood of failure

Solution

- Prepared large volume of grid characteristic data
- Evaluated 100s of modeling approaches
- Compared results with confusion matrices
 - Transparency of model performance
 - Easily validated what had been built

Impact

- Able to quickly identify areas in need of repair
- Able to fix performance degradation before it occurs
- Optimized allocation of repair budget to largest need
- Built a functioning 'monetized risk approach'
 - Able to calculate ROI of risk mitigation activities
 - Used to actively manage a risk program
 - Can choose which risk mitigation activities to fund

Problem type: Predictive maintenance

Universal relevance: Predictive maintenance is a machine learning application that is widely understood in manufacturing, but has far-reaching implications for many other industries. Every business has high-cost assets that will eventually break-down.

Final Thoughts

Many businesses wait until their data is perfect, or until they have a 'grand-slam' use case before they experiment with AI. This approach is detrimental and anti-progressive. Data science is just like any other scientific endeavor. You must test and iterate, in this case with a portfolio of potential use cases.

You may not diagnose diseases, mine rock or operate Crack Gas Compressors, but you can still draw inspiration from the way these organizations are using AI to drive impact. We hope you've found this library of use cases helpful to spark your imagination and open your eyes to new potential ways to leverage AI and Machine Learning.

If you would like more help, RapidMiner will guide you through a personalized AI Assessment to identify and prioritize the right portfolio of use cases for your business.

[Request an AI Assessment today.](#)

