DZone

THE 2018 DZONE GUIDE TO

Artificial Intelligence

AUTOMATING DECISION-MAKING

VOLUME II

BROUGHT TO YOU IN PARTNERSHIP WITH

outsystems

Dear Reader,

I would like to offer a warm welcome to DZone's 2018 Guide to Artificial Intelligence. Artificial intelligence (AI) and all of its subgroupings have exploded this year. You can hardly find a company not talking about their deep learning, AI, or machine learning (ML) strategies. I have done a number of talks at major conferences in Europe and North America recently, and have noticed that any talk mentioning anything related to AI is well-attended. People are constantly coming up to me asking how they can get involved with AI at their company.

In this Guide, we will focus on self-running systems with AI, using machine learning, advanced predictive analytics, natural language processing, AI/ML open-source projects, and Industrial IoT (IIoT). As you will see, AI is improving many existing technologies and solving new problems. One of the first things you will notice is the breadth of use cases that AI can help teams solve. This Guide shows a variety of interesting ways to utilize AI from credit card fraud detection, to predictive analytics, to IIoT. You will see how important a tool that AI has become.

It will be no surprise to anyone in enterprise software to learn that there is a great deal of innovation being done withartificial intelligence in the open-source community. Often, the research done in universities and the largest internet companies is released as open-source frameworks such as TensorFlow, Apache MXNet, and Pytorch. That research is also leading to the release of algorithms and pre-trained machine learning and deep learning models to the open-source community. This openness is enabling universities, companies, and technologists to collaborate at scale on some of the most difficult problems out there. We are seeing great advancements in computer vision, natural language processing, voice recognition, sentiment analysis, and fraud detection.

Thank you for downloading this Guide. As the world of AI, machine learning, and deep learning rapidly improves and expands, check back at DZone to catch up with the latest in AI trends, frameworks, news, tutorials, and technologies. As with any great community, your feedback matters, and we would love to hear it. As much as AI has grown, there is much work to be done. AI without the context of other software systems, big data, integration, and real-world examples becomes just noise. Check out DZone.com daily to stay in touch with the latest in this ever-changing area.

Happy reading!

By Tim Spann

SENIOR SOFTWARE ENGINEER AT HORTONWORKS & DZONE ZONE LEADER

Table of Contents

DZone is...

BUSINESS & PRODUCT

MATT TORMOLLEN
CEO

MATT SCHMIDT
PRESIDENT

JESSE DAVIS
EVP, TECHNOLOGY

KELLET ATKINSON
MEDIA PRODUCT MANAGER

PRODUCTION

CHRIS SMITH
DIRECTOR OF PRODUCTION

ANDRE POWELL
SR. PRODUCTION COORD.

ASHLEY SLATE
DESIGN DIRECTOR

G. RYAN SPAIN
PRODUCTION COORD.

BILLY DAVIS
PRODUCTION COORD.

NAOMI KROMER
SR. ACCOUNT MANAGER

JASON BUDDAY
ACCOUNT MANAGER

MICHAELA LICARI
ACCOUNT MANAGER

EDITORIAL

CAITLIN CANDELMO
DIR. OF CONTENT & COMMUNITY

MATT WERNER
PUBLICATIONS COORDINATOR

SARAH DAVIS
PUBLICATIONS ASSOCIATE

MICHAEL THARRINGTON
CONTENT & COMMUNITY MANAGER II

KARA PHELPS
CONTENT & COMMUNITY MANAGER

TOM SMITH
RESEARCH ANALYST

MIKE GATES
CONTENT TEAM LEAD

JORDAN BAKER
CONTENT COORDINATOR

ANNE MARIE GLEN
CONTENT COORDINATOR

ANDRE LEE-MOYE
CONTENT COORDINATOR

LAUREN FERRELL
CONTENT COORDINATOR

LINDSAY SMITH
CONTENT COORDINATOR

SALES

CHRIS BRUMFIELD
SALES MANAGER

FERAS ABDEL
SALES MANAGER

ALEX CRAFTS
DIRECTOR OF MAJOR ACCOUNTS

JIM HOWARD
SR. ACCOUNT EXECUTIVE

JIM DYER
SR. ACCOUNT EXECUTIVE

ANDREW BARKER
SR. ACCOUNT EXECUTIVE

BRIAN ANDERSON
ACCOUNT EXECUTIVE

SARA CORBIN
ACCOUNT EXECUTIVE

BRETT SAYRE
ACCOUNT EXECUTIVE

MARKETING

SUSAN WALL
CMO

AARON TULL
DIRECTOR OF MARKETING

LAUREN CURATOLA
MARKETING SPECIALIST

KRISTEN PAGÀN
MARKETING SPECIALIST

JULIAN MORRIS
MARKETING ASSOCIATE

Executive Summary

BY MATT WERNER - PUBLICATIONS COORDINATOR, DZONE

Last year's *Guide to Artificial Intelligence* was one of the most popular Guides DZone released in 2017, and it's been clear from our coverage of big data, web development, performance monitoring, and the Internet of Things that AI is making waves across the entire development world. This year, however, we expanded our audience survey, and saw several declines in those implementing AI libraries or methods in their applications and systems. With such a stark change within one year, we've analyzed this data and included several helpful articles from AI professionals to help realize the potential of artificial intelligence and machine learning.

CHANGES FROM 2017

Data: Between 2017 and 2018, there was a 10% decrease in those implementing AI or machine learning in their personal projects, and a 2% decrease in those implementing them for work-related projects. When asked what prevents them from being interested in AI/ML, there was a 22% decrease in those who mentioned time being a factor between 2017 (48%) and 2018 (26%), but a 22% increase in those who did not see an organizational benefit or use case between 2017 (28%) and 2018 (50%). There was a small change in the number of people who saw a lack of experience as a challenge between 2017 (40%) and 2018 (38%).

Implications: While a lack of developer experience is still an issue in adopting AI, more developers seem to have ignored any hype around AI systems and do not see an organizational benefit or use case for their applications.

Recommendations: The perceived lack of use cases could be because other tasks are being prioritized over implementing AI, or the time it would take to implement AI into legacy systems is not considered to be worth it. Analyze the time and money it would take to seriously implement AI/ML capabilities, as well as user stories that would be improved by these features.

ORGANIZATIONS BELIEVE IN AI

Data: 3% more organizations are interested in and have invested in AI/ML than last year (28% vs. 31%), and organizations interested in AI/ML, but not invested, grew slightly to 33% from 31% last year. 37% of organizations do not see an organizational benefit to AI, compared to 61% last year.

Implications: While individual developers are not as interested in AI now, their organizations and businesses still see the benefit and are working to unlock the potential in their systems.

Recommendations: In 2018, 45% of organizations cite developer experience as a major challenge in adopting and investing in AI. Regardless of developers' personal feelings, organizations are becoming more and more interested in implementing AI/ML capabilities in their applications — and learning how to utilize common frameworks and libraries such as TensorFlow will be useful for developers looking to advance their careers.

HOW DEVELOPERS USE MACHINE LEARNING

Data: For those who have implemented AI/ML in their personal projects, the most common use case has been prediction (24%), followed by classification (19%), detection (15%), automation (14%), and recommendations (14%).

Implications: The most common use case for machine learning is to assist in predictive analytics and extrapolate future data from current trends. This is very difficult to do, as past performance does not necessarily indicate future results, but can be useful for projections. Classifying pieces of data is also very popular, especially for sorting and making sense of unstructured data.

Recommendations: All of the listed applications of machine learning can help save time and give users and developers much-needed insights or features. Use almost any major website and you can see these in action, such as recommended items on Amazon or recommended titles on Netflix. Take the time to learn some basics about statistics and how to properly analyze data, and think about what other metrics you could use and how AI/ML can help you measure and predict them.

Key Research Findings

BY JORDAN BAKER - CONTENT COORDINATOR, DZONE

DEMOGRAPHICS

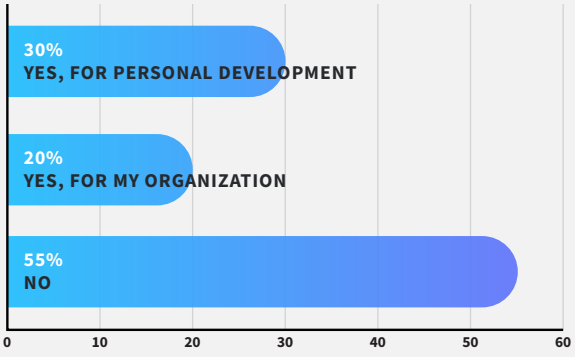
403 respondents from DZone's community of developers, dev leads, architects, and technologists took the 2018 DZone AI/ Machine Learning Survey. The demographics for this group break down as given below:

- 38% identity as software developers/engineers, 16% as a developer team leads, and 15% as architects.
- The average respondent has 13 years of experience in the software industry.
- 30% live in South Central Asia, 25% in the USA, and 21% in Europe.
- 29% work for organizations that employ 100-999 employees; 19% work for organizations that house 10,000+ employees; 18% work for organizations with 1,000-9,999 employees.
- 18% work in the software vendor industry, 17% in finance/banking, and 11% in consulting.

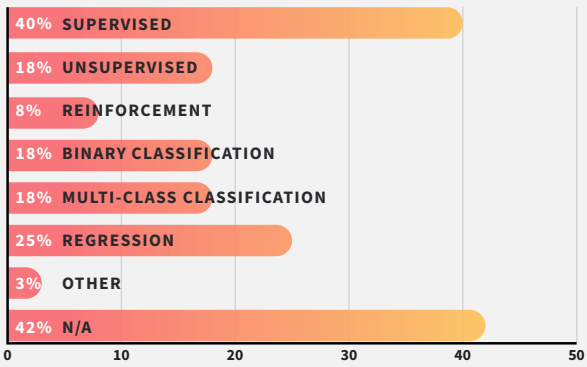
AI FOR PERSONAL VS. PROFESSIONAL PROJECTS

Given the complexity of AI, and the fact that large, enterprise-level companies like Google, Netflix, and Amazon all back AI and machine learning projects, one would expect the majority of developers who use AI to do so in the work place. And, as 286 out of the 403 respondents to this survey work for organizations that employ over 100 people (i.e. enterprise-level organizations), one would, again, expect the data to slant toward AI for professional projects. But, in fact, 30% of respondents told us they have used AI in personal projects, whereas 20% have implemented it in professional projects. Something that might be skewing these numbers is that 55% of respondents have not implemented AI/ML in a project before. But of the 338 respondents who told us they have not implemented AI in their projects, 20% have at least toyed around with one or more of the following types of machine learning: supervised machine learning, regression, unsupervised machine learning, binary classification, and/or multi-class classification.

GRAPH 01. Have you ever implemented or used AI or machine learning in systems you developed?



GRAPH 02. Which of these types of machine learning have you used?



The five machine learning techniques listed above also turned out to be the most popular ways of working with machine learning among our respondents. 40% of survey takers reported having worked with supervised machine learning and 25% have used regression algorithms. The final three listed — unsupervised machine learning, binary classification, and multi-class classification — all received 18% of the vote in this category. Interestingly, the percentages of respondents who've worked with these types of machine learning both at home and the office are more or less identical. For example, 30% of respondents have used supervised machine learning in systems developed for work, whereas 33% have used supervised machine learning for personal projects. The only real deviation seems to be in the use of multi-class classification. Whereas 17% of those who use AI/machine learning at work have used multi-class classification, 13% of have used this type of machine learning for personal projects. While this is not a large difference, it does, nonetheless, represent the biggest percentage gap in types of machine learning used on personal vs. professional projects.

Looking at how all respondents use AI and machine learning, regardless of personal or professional projects, the largest use case by far is for prediction. 43% of those who took the survey told us that they are using machine learning for its abilities with predictive analytics. The second most popular purpose for machine learning was classification, totaling 31% of respondents. While not nearly as popular as prediction and classification, some other popular machine learning use cases were recommendations (23%), detection (23%), and optimization (21%).

Given that 55% of respondents told us they have never implemented AI or machine learning in the systems they've developed, let's quickly explore why some developers are hesitant to adopt AI. When we asked, "What issues prevent you from being interested in AI/machine learning?", 50% of the 221 respondents to this question answered that there are not enough use cases for applications built on this technology. Another 38% said there's not enough developer experience in the area, and 26% said time played a factor.

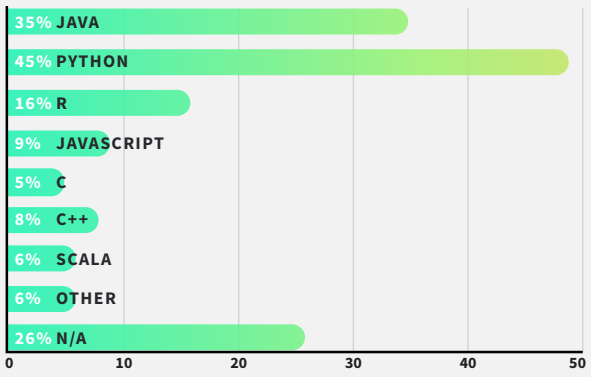
AI IN THE INDUSTRY

Though more respondents to this survey have used AI in personal projects than in professional development projects, AI/machine learning is quickly gaining steam in the industry. 33% of survey respondents reported that their organization is interested in AI or machine learning, but has not yet invested, and an additional 31% of respondents work for organizations that are both actively invested and interested in AI. In fact, only 16% of survey takers told us their organization is not interested in AI/machine learning.

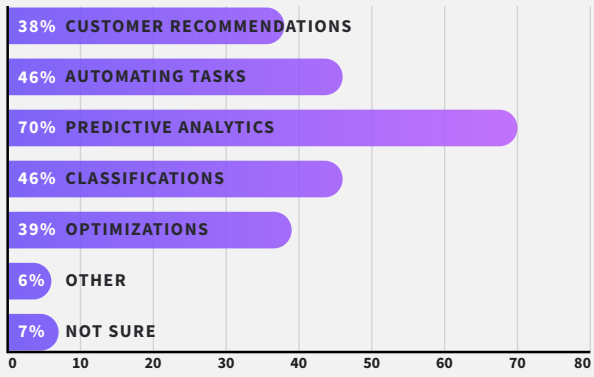
When we look at the factors that contribute to this high percentage of organizations exploring AI/machine learning, the organizational-level use cases for this technology nearly mirror the uses for AI that attract individual developers. 70% of respondents reported that their organization is using AI/machine learning for predictive analytics — far and away the largest use case reported in this survey. Other popular organizational uses for AI/machine learning included automating tasks (47%), classification (46%), optimization (40%), and customer recommendations (38%). Given that 75% of respondents are currently developing web applications, it seems a fair assumption to say that most of these types of machine learning are being used to serve up a better experience (whatever that may mean for a particular application) to users.

Due to this wide interest that organizations are taking in AI and machine learning, there's also a concerted effort being put forth to train developers to use this technology. When we

GRAPH 03. Which languages do you use for machine learning development?



GRAPH 04. What is your organization trying to achieve with AI/machine learning?



asked respondents what their organization is doing to pursue AI/machine learning, 56% reported an investment in developer training. Other popular answers were educating upper management on AI (35%) and utilizing open data initiatives and datasets (32%).

LANGUAGES, LIBRARIES, AND APIS

In the 2017 DZone Guide to Artificial Intelligence, 41% of respondents reported using Java and 40% reported using Python for machine learning development. In this year's survey, Java and Python traded places. 45% of survey takers told us they use Python for machine learning development, and 35% reported using Java. The third most popular ML language in this year's survey, R, stayed stagnant year-over-year, coming in at 16%. When we look at these three languages (Python, Java, and R) and how their use varies between personal and professional projects, we see some interesting fluctuations in the percentages, as well. Among those respondents who have developed AI/machine learning systems for work, 37% use Python, 24% use Java, and 13% use R. Among the hobbyist AI/ML developers, 40% are using Python, 26% are using Java, and 9% are using R.

When it comes to widely usable machine learning libraries, TensorFlow has set the bar. This open-source library, originally created by Google engineers, has over 105,000 stars on GitHub, making it the most popular option among developers for AI/ML today (source). The popularity of TensorFlow among the wider developer community was mirrored in the DZone community, with 27% of survey respondents for this Guide reporting to use the TensorFlow library in their projects. Other popular open-source libraries for AI/ML among the developer community at large are Sckit-Learn, which has over 29,000 starts on GitHub [source], and Spark MLLib, which has 18,000+ stars (source). Again, the popularity of these libraries was reflected in the DZone AI survey. 15% of survey respondents have worked with Sckit-Learn (the second most popular option among survey takers) and 14% have worked with Spark MLLib. Another popular option among our survey respondents, though not open-source, was the Amazon Machine Learning service.

Given the popularity of the above libraries and services (and the open-source nature of many of them), it is no surprise that 40% of respondents generally use APIs when developing machine learning applications.

POPULAR ALGORITHMS AND THEIR USES

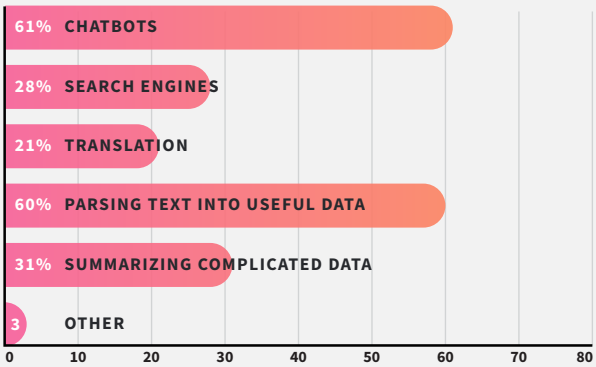
When asked which algorithms they use for machine learning, respondents gave four main answers: linear regression, decision trees, neural networks, and natural language processing (NLP). Linear regression received the most interest in this survey, with 37% of respondents using this algorithm. Decision trees were a close second at 35%, neural networks were third with 31%, and NLP received 29% of respondents' votes.

Though NLP did not rank as high as expected in our algorithms poll, natural language processing and neural networks are two of the more popular algorithms in the fields of AI and machine learning. Thus, we asked respondents how they are using these two important algorithms.

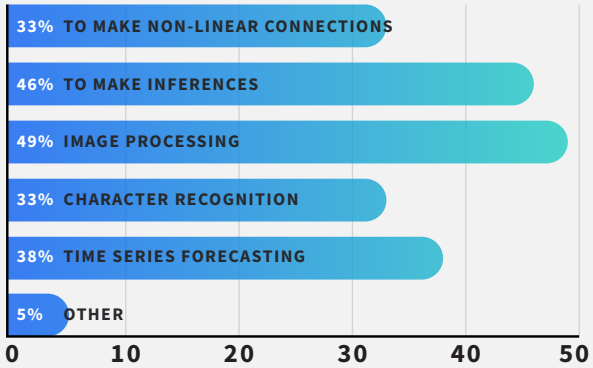
50% of respondents who use neural networks in their projects use them for image processing. The second most popular use for neural nets, with a 45% adoption rate among survey takers, was to make inferences. Some other popular uses of this algorithm were time series forecasting (38%), making non-linear connections (33%), and character recognition (33%).

For the respondents who use NLP, 61% utilize this algorithm in their chatbots and 60% for parsing text into useful data. These were by far the most popular uses for NLP in this year's survey. This third most popular option was for summarizing complicated data (31%) and for use in search engines (28%).

GRAPH 05. What are you using natural language processing for?



GRAPH 06. What are you using neural networks for?



Using Apache Ignite's Machine Learning for Fraud Detection at Scale

QUICK VIEW

01. Get a problem statement regarding credit card fraud and the need to use modern machine learning tools for analysis and prevention.

02. Get an outline of an approach to perform the model building and analysis using the proposed dataset.

03. Get a description of the Apache Ignite algorithm to tackle the problem, develop code, run the code, and present the output.

BY AKMAL CHAUDHRI
TECHNOLOGY EVANGELIST AT GRIDGAIN

Apache Ignite is a very versatile product that supports a wide range of integrated components. These components include a machine learning (ML) library that supports popular ML algorithms, such as linear regression, k-NN classification, and k-means clustering.

The ML capabilities of Ignite provide a wide range of benefits. For example, Ignite can work on the data in-place, avoiding costly ETL between different systems. Ignite can also provide distributed computing that includes both the storage and manipulation of data. The ML algorithms implemented in Ignite are optimized for distributed computing and can use Ignite's collocated processing to great advantage. Ignite can also act as a sink for streaming data, allowing ML to be applied in real-time. Finally, ML is often an iterative process, and context may change while an algorithm is running. Therefore, to avoid loss of work and delay, Ignite supports partition-based datasets, which makes it tolerant to node failures.

Ignite ships with code examples that demonstrate the use of various ML algorithms on some well-known datasets. These code examples can work standalone, making it very easy to get started with the ML library, and the examples can be used as templates for user-specific problems.

With these benefits in mind, let's undertake some exploratory analysis and write some code for a real-world problem. Specifically, let's look at how Ignite could help with credit card fraud detection.

BACKGROUND

Today, credit card fraud is still a major problem experienced by many financial institutions. Historically, checking financial transactions has been a manual process. However, we can now apply ML techniques to identify potentially fraudulent transactions and, therefore, be able to develop real-time fraud detection systems that can act much faster and help stop these fraudulent transactions in their tracks.

THE DATASET

A suitable dataset for credit card fraud detection is available through Kaggle [1], provided by the Machine Learning Group at Université Libre de Bruxelles (ULB). The data are anonymized credit card transactions that contain both genuine and fraudulent cases. The transactions occurred over two days during September 2013 and the dataset contains a total of 284,807 transactions of which 492 are fraudulent, representing just 0.172% of the total. This dataset, therefore, presents some challenges for analysis as it is highly unbalanced.

The dataset consists of the following fields:

- **Time**: The number of seconds elapsed between a transaction and the first transaction in the dataset.
- **V1 to V28**: Details not available due to confidentiality reasons.
- **Amount**: The monetary value of the transaction.
- **Class**: The response variable (0 = no fraud, 1 = fraud).

MACHINE LEARNING ALGORITHM

According to Andrea Dal Pozzolo, who was involved in the collection of the original dataset, fraud detection is a classification problem [2]. Also, since investigators may only review a limited number of transactions, the probability that a transaction is fraudulent is more important than the true classification. Therefore, a good algorithm to use for the initial analysis is logistic regression. This is because the outcome has only two possible values, and we are also interested in the probability.

DATA PREPARATION

As previously mentioned, the dataset is highly unbalanced. There are a number of solutions we can use to manage an unbalanced dataset [3]. The initial approach we can take is to under-sample. We will keep all

the 492 fraudulent transactions and reduce the number of non-fraudulent transactions. There are several ways we could perform this dataset reduction:

1. Randomly remove majority class examples.
2. Select every nth row from the majority class examples.

For our initial analysis, let's use the second approach and select every 100th majority class example.

We know that there are columns V1 to V28, but not what these represent. The Amount column varies significantly, between 0 and 25691.16. We can remove the Time column, since it does not provide the actual time of a transaction. There are no missing values in the dataset.

Another decision that we need to make is whether to normalize the data. For our initial analysis, we won't use normalization.

One approach to data preparation for this credit card fraud problem is described by Kevin Jacobs [4]. Simple analysis can often provide good initial insights and help refine the strategy for further data analysis. Using the approach described by Jacobs, we'll create our training and test data using scikit-learn. We'll then load the data into Ignite storage and perform our logistic regression using Ignite's Machine Learning Library.

Once our training and test data are ready, we can start coding the application. You can download the code from GitHub [5] if you would like to follow along. We need to do the following:

1. Read the training data and test data
2. Store the training data and test data in Ignite
3. Use the training data to fit the logistic regression model
4. Apply the model to the test data
5. Determine the confusion matrix and the accuracy of the model

READ THE TRAINING DATA AND TEST DATA

We have two CSV files with 30 columns, as follows:

1. V1 to V28
2. Amount
3. Class (0 = no fraud, 1 = fraud)

We can use the following code to read-in values from the CSV files:

```
private static void loadData(String fileName,
IgniteCache<Integer, FraudObservation> cache)
        throws FileNotFoundException {
   Scanner scanner = new Scanner(new File(fileName));
   int cnt = 0;
   while (scanner.hasNextLine()) {
      String row = scanner.nextLine();
      String[] cells = row.split(",");
```

CODE CONTINUED ON NEXT COLUMN

```
      double[] features = new double[cells.length - 1];

      for (int i = 0; i < cells.length - 1; i++)
         features[i] = Double.valueOf(cells[i]);
      double fraudClass = Double.valueOf(cells[cells.
length - 1]);

      cache.put(cnt++, new FraudObservation(features,
fraudClass));
   }
}
```

The code reads the data line-by-line and splits fields on a line by the CSV field separator. Each field value is then converted to double format and then the data are stored in Ignite.

STORE THE TRAINING DATA AND TEST DATA IN IGNITE

The previous code stores data values in Ignite. To use this code, we need to create the Ignite storage first, as follows:

```
IgniteCache<Integer, FraudObservation> trainData =
getCache(ignite, "FRAUD_TRAIN");
IgniteCache<Integer, FraudObservation> testData =
getCache(ignite, "FRAUD_TEST");
loadData("src/main/resources/resources/fraud-train.
csv", trainData);
loadData("src/main/resources/resources/fraud-test.
csv", testData);
```

The code for `getCache()` is implemented like so:

```
private static IgniteCache<Integer, FraudObservation>
 getCache(Ignite ignite, String cacheName) {
   CacheConfiguration<Integer, FraudObservation>
cacheConfiguration = new CacheConfiguration<>();
   cacheConfiguration.setName(cacheName);
   cacheConfiguration.setAffinity(new
RendezvousAffinityFunction(false, 10));
   IgniteCache<Integer, FraudObservation> cache =
ignite.createCache(cacheConfiguration);
   return cache;
}
```

USE THE TRAINING DATA TO FIT THE LOGISTIC REGRESSION MODEL

Now that our data are stored, we can create the trainer, as follows:

```
LogisticRegressionSGDTrainer<?> trainer = new
LogisticRegressionSGDTrainer<>(new UpdatesStrategy<>(
   new SimpleGDUpdateCalculator(0.2),
   SimpleGDParameterUpdate::sumLocal,
   SimpleGDParameterUpdate::avg
), 100000, 10, 100, 123L);
```

We are using Ignite's Logistic Regression Trainer with stochastic gradient descent (SGD). The learning rate is set to 0.2 and controls how much the model changes. We have also specified the maximum number of iterations as 100,000 and the seed as 123.

We can now fit the logistic regression model to the training data, as follows:

```
LogisticRegressionModel mdl = trainer.fit(
        ignite,
        trainData,
        (k, v) -> v.getFeatures(),     // Feature
extractor.
        (k, v) -> v.getFraudClass()    // Label
extractor.
).withRawLabels(true);
```

Ignite stores data in a key-value (K-V) format, so the above code uses the value part. The target value is the fraud class and the features are in the other columns.

APPLY THE MODEL TO THE TEST DATA

We are now ready to check the test data against the trained logistic regression model. The following code will do this for us:

```
int amountOfErrors = 0;
int totalAmount = 0;
int[][] confusionMtx = {{0, 0}, {0, 0}};
try (QueryCursor<Cache.Entry<Integer,
FraudObservation>> cursor = testData.query(new
ScanQuery<>())) {
   for (Cache.Entry<Integer, FraudObservation>
testEntry : cursor) {
      FraudObservation observation = testEntry.
getValue();
      double groundTruth = observation.
getFraudClass();
      double prediction = mdl.apply(new
DenseLocalOnHeapVector(observation.getFeatures()));
      totalAmount++;
      if ((int)groundTruth != (int)prediction)
         amountOfErrors++;
      int idx1 = (int)prediction;
      int idx2 = (int)groundTruth;
      confusionMtx[idx1][idx2]++;
      System.out.printf(">>> | %.4f\t | %.0f\t\t\
t|\n", prediction, groundTruth);
   }
}
```

DETERMINE THE CONFUSION MATRIX AND THE ACCURACY OF THE MODEL

Now we can check by comparing how the model classifies against the actual fraud values (ground truth) using our test data.

Running the code gives us the following summary:

```
>>> Absolute amount of errors 80
>>> Accuracy 0.9520
>>> Precision 0.9479
>>> Recall 0.9986
>>> Confusion matrix is [[1420, 78], [2, 168]]
```

For the confusion matrix, we have the following:

	NO FRAUD	FRAUD
NO FRAUD	1420	78 (Type I error)
FRAUD	2 (Type II error)	168

SUMMARY

Our initial results look promising, but there is room for improvement. We made a number of choices and assumptions for our initial analysis. Our next steps would be to go back and evaluate these to determine what changes we can make to tune our classifier. If we plan to use this classifier for a real-time credit card fraud detection system, we want to ensure that we can catch all the fraudulent transactions and also keep our customers happy by correctly identifying non-fraudulent transactions.

Once we have a good classifier, we can use it directly with transactions arriving into Ignite in real-time. Additionally, with Ignite's continuous learning capabilities, we can refine and tune our classifier further with new data, as the data arrive.

Finally, using Ignite as the basis for a real-time fraud detection system enables us to obtain many advantages, such as the ability to scale ML processing beyond a single node, the storage and manipulation of massive quantities of data, and zero ETL.

REFERENCES

[1] kaggle.com/mlg-ulb/creditcardfraud

[2] slideshare.net/dalpozz/adaptive-machine-learning-for-credit-card-fraud-detection

[3] analyticsvidhya.com/blog/2017/03/imbalanced-classification-problem

[4] data-blogger.com/2017/06/15/fraud-detection-a-simple-machine-learning-approach

[5] https://github.com/VeryFatBoy/ignite-ml-examples/

AKMAL CHAUDHRI is a Technology Evangelist at GridGain. His role is to help build the global Apache Ignite community and raise awareness through presentations and technical writing. He has published and presented widely and edited or co-edited ten books. He holds a BSc (1st Class Hons.) in Computing and Information Systems, MSc in Business Systems Analysis and Design, and a PhD in Computer Science. He is a Member of the British Computer Society (MBCS) and a Chartered IT Professional (CITP).

How AI Is Bringing Crucial Machine Intelligence to IoT

BY CATE LAWRENCE
MARKETING ASSOCIATE, MOMENTA PARTNERS AND DZONE ZONE LEADER

QUICK VIEW

01. AI is moving IoT beyond collecting deluges of data to making sense of data.

02. We are witnessing the growth of AI-powered analytics platforms for the enterprise market and the enablement of predictive and prescriptive analytics and adaptive/continuous analytics.

03. The combination of AI and IoT is disrupting industries such as manufacturing and agriculture.

04. AI is finally making connected devices not just connected but smart.

We find ourselves in a pivotal era in technology, where the things we've been promised for years like augmented reality, advanced 3D printing, artificial intelligence, and autonomous vehicles are becoming a reality (albeit not necessarily in the forms we'd like them yet, or in all locations or price points). As technology evolves, technologists across disciplines and industries are sharing knowledge and research, leading to even more discoveries and cross-practice convergence.

At the nexus of this evolution is the Internet of Things. Its ubiquity may receive the most attention in consumer and residential products, but the validation of the ability to change things for the better is evidenced most strongly in Connected Industry. But, smart consumer devices are just starting to catch up. Add artificial intelligence to the mix and things really start getting interesting:

THE BENEFITS OF ARTIFICIAL INTELLIGENCE

It's easy to see AI as something of a party trick, (like the dog/muffin pic), a killer robot (not such a great party as we thought), or a buzzword that rolls off the tongue of a startup CEO like something from a checklist. In reality, artificial intelligence has achieved a wide range of capabilities from image recognition to language processing and data analysis. It enables machines to make decisions and deductive reasoning faster than humans. It creates an extra layer on what IoT can achieve, almost an inter-dependence, such as in the growth of AI-powered analytics platforms for the enterprise market and the enablement of predictive and prescriptive analytics and adaptive/continuous analytics.

The sheer deluge of data emanating out of connected devices can add little value to anyone is it is simply stored and collected without rigorous analysis. AI-powered analytics enable the division between time-sensitive data that is processed at the edge (such as that of a piece of connected safety equipment) and more extraneous data that can be processed in larger volume and with less urgency in the cloud.

Each year, what it makes possible succeeds that of the proceeding ten years. We can expect that AI will lay the foundation for an acceleration in innovation over the next few years, boosting some sectors of the economy and completely restricting some industries. While programmers still control the capabilities of AI at present, this may not be the case forever. Let's take a look at some of the key use cases where the convergence of IoT and AI are leading to ever bigger resolutions than the technologies alone could ever achieve.

PRECISION AGTECH

The world of agricultural technology, or agtech, is rapidly evolving. It's automating laborious tasks from seed sowing to crop picking, and picking up the slack in roles where farmers are struggling to recruit seasonal staff. It's also providing farmers and growers with greater knowledge and insight into their crops and livestock than ever before.

Notably, around 20% of the world's food production is grown within cities rather than rural areas, and inherent in this is the multi-billion dollar industry of indoor growing and hydroponics. The industry includes $5 billion in urban farming in the US and $5.7 billion for legal cannabis production.

Today, farms can leverage IoT to remotely monitor soil moisture, crop growth, smart connected harvesters, and irrigation equipment. Then farmers can analyze operational data combined with third-party information, such as weather services, to provide new insights and improve decision-making.

Farmers and growers who have long relied on almanacs for historical references of their crops can now enjoy the ease of a tablet or mobile phone. For example, at the Pago Aylés winery, in collaboration with measurement agricultural scientists from remOT Technologies, has invested in an IoT project with Libelium technology to gain efficient production and a predictive vineyard management model. Strategically placed sensors were installed to measure:

- Temperature, humidity, and environmental pressure.
- Soil temperature.
- Ground humidity.
- Rainfall, wind speed, and direction.

Winemakers are then able to access the data on their smartphones at any time through a corresponding app which, when combined with historical data, enables growers to establish patterns and predictive models on the behavior of the vineyard. The use of AI means faster data collection and processing and the data collected means they are not only able to be more precise about watering, but also to accurately predict their yield and anticipate the demand before forecasts of purchases or sales according to the estimated production.

PREDICTIVE MAINTENANCE IN MANUFACTURING

One of the most compelling aspects of industrial IoT is the ability of sensor technology to solve problems that have plagued traditional industries for years, if not decades. One such problem is machine maintenance and repairs. Up until recently, such maintenance was typically time-consuming, costly, and limited by the challenge of finding appropriately skilled workers. A broken machine could take hours or even days to fix (especially if parts needed to be sourced), and the downtime could result in as high as millions of dollars worth of sellable products.

Machine maintenance is traditionally a laborious process that requires sending a physical person around each factory/plant/workplace to inspect individual machines, typically on a set schedule. A combination of sensors and AI offer a more effective solution. One example is the work of 3DSignals. They utilize sensor tech to monitor machines through sound. Their system can extend to a range of machines "based on the knowledge of how similar machines are supposed to sound and also learning the very specific sound acoustics of specific machines." This all leads to increased efficiency in maintenance and the ability to predict problems so that an engineer can respond as needed rather than only within a pre-existing preventive maintenance time frame.

3DSignals' acoustic monitoring and deep learning technology monitors sensory data from production line machinery, identifies anomalies, classifies patterns of equipment failure, and predicts issues before they interrupt production. This reduces downtime and saves significant time and money in lost production.

YOUR PERSONAL IOT IS *FINALLY* GETTING SMART, THANKS TO AI

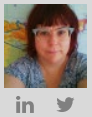
What if there was a way to make connected products personal, instinctual, and truly targeted to the individual? Start-up Neura has created an AI engine that turns IoT environments into connected universes that allow companies to connect with their customers during the most meaningful moments. Neura enables consumers to take control of their smart home devices — Amazon Echo, Nest thermostat, Hue Lights, Ring Smart doorbell, refrigerators, and more — and make their smart homes more intelligent with the integration of true AI. Their AI engine integrates with multiple data channels to provide situational awareness for individual customers, ensuring personalized, highly relevant engagement.

Neura's tech relies on the reality that the users of IoT tech are surrounded by connected devices throughout most of their day, whether WiFi at work, a smartwatch, a Bluetooth headset, or their connected car. Because of this, an endless flow of connections with their specific patterns and clusters is exposed through predefined moments. Neura's cognitive computing techniques find the patterns that define the users' daily lives to detect and understand the relations between users and their surroundings, their work and home routine, exercise habits, etc. and provide companies with a simple to consume API.

The Neura SDK is integrated into an app and that starts pulling the data from the phone's sensors as well as WiFi and Bluetooth signals. This then feeds into a hybrid AI engine. Therefore, the majority is cloud-based and a lightweight version is in the SDK. Your products access specific insights and predictions via API calls like that you've gone running or left the house. The lightweight AI engine is able to react in real time to an anomaly and change the consequential actions. For example, if you usually go to bed at 10 PM and tonight you're awake, drinking in a bar, the things that usually happen (thermostat change, door locking) aren't going to occur because you're not where you're usually are.

It's easy to be cynical about AI and to question its validity as more than a marketing buzzword. But the reality is that its use, when combined with other technologies such as IoT, is making ever greater advancements possible. Its evolution is coming at us fast and hard and we'd better stay alert if we want to be part of the decision making. we'd better keep an eye out if we want to keep up.

CATE LAWRENCE Is an experienced technology journalist and writer based in Berlin. Cate brings an understanding and focus on IoT, biohacking, and future technology to DZone. She likes to look behind the technology to see the challenges and opportunities its creation poses for developers, users, law makers, and society.



AI and Low-Code: The Ultimate BFFs

After the initial buzz around AI in the 90s, it took a long winter nap but has now come back with a vengeance. And with good reason:

1. The widespread availability of cheap storage and proliferation of digital channels means organizations are capturing more data than ever before.
2. Sophisticated cognitive services can now be accessed from the cloud on inexpensive computing devices, meaning artificial intelligence is available to anyone.
3. The low-code development platform market is booming, predicted to grow from USD 4.32 billion in 2017 to USD 27.23 billion by 2022, at a CAGR of 44.5%.

As organizations gather petabytes of data with potentially valuable customer insights, what can they do to unlock the secrets held within? Enter AI algorithms to feed on the huge data sets to find patterns. But what's missing is a way to harness the power of AI. Low-code platforms are the perfect glue, enabling developers to quickly connect data to technology and build new solutions.

WHY OUTSYSTEMS?

- Visual, model-driven development means "intelligent" applications can be built without expensive, hard-to find resources.
- Rapid experimentation and built-in feedback enable organizations to test and validate new concepts.
- Pre-built integrations with leading data sources and AI cloud services accelerate solution development.

Organizations taking advantage of this combination can disrupt established industries, breaking new ground while bringing more value to customers.

WRITTEN BY MIKE HUGHES
SENIOR DIRECTOR, PRODUCT MARKETING, OUTSYSTEMS

PARTNER SPOTLIGHT

OutSystems Low-Code Development Platform

OutSystems is the #1 low-code platform for digital transformation - build mobile apps, web portals, mission-critical systems, and more.

outsystems

CATEGORY
High productivity, low-code application development and delivery platforms.

RELEASE SCHEDULE
Annual major releases and monthly feature releases

OPEN SOURCE?
No, but we offer a free personal edition

CASE STUDY
Professional services firm Deloitte LLP used the OutSystems platform to build an innovative AI and machine-learning app. The app helps customers streamline the process of screening for "conduct risk." Called BEAT (Behavioral and Emotional Analytics Tool), the app uses technology to spot potentially risky sales interactions rather than requiring customers to manually review hundreds of thousands of transactions. Deloitte used OutSystems to create a proof of concept in under a week and deliver a fully functioning app in five months. "Leveraging the power of AI in our BEAT application has massively reduced both the time and cost of reviewing call records. This is a game-changer," said Rui Vaz, partner with Deloitte.

STRENGTHS
1. Visual, low-code full-stack web and mobile application development
2. Integration with everything: Pre-built connectors to leading AI/ML platforms
3. Change tracking and automatic change impact analysis
4. One-click deployment and continuous delivery for even the most complex apps
5. High availability, scalability, and enterprise-grade security for running apps

NOTABLE USERS
- Toyota
- Ricoh
- GM Financial
- Chevron
- Schneider Electric

WEBSITE outsystems.com **TWITTER** @OutSystems **BLOG** outsystems.com/blog

Pushing AI to the Edge: Be Part of the Revolution

BY WOLF RUZICKA
CHAIRMAN, EASTBANC TECHNOLOGIES

QUICK VIEW

01. Every business has a target audience who cares about their experience so the potential for AI to have an impact in your business is huge.

02. Edge AI pushes you to think bigger than your current parameters, competitors, or even your market.

03. Start small. We've seen some of the most valuable business insights derived from surprisingly small data sets.

Artificial intelligence (AI) surrounds us. It unlocks our phones, creates our shopping list, navigates our commute, and cleans spam from our email. It's making customers' lives easier and more convenient. Once you've experienced AI in action, it's difficult to go back. With edge computing becoming a thing, AI on-the-edge is following suit.

Right now, edge AI seems like it's only a game for the most cutting-edge companies (Apple, Amazon, Tesla, to name a few). Traditional enterprises aren't really embracing it. The thinking: The technology isn't mature enough. It will cost too much. There's no proven ROI. Our competition isn't doing it yet, so why bother now.

But waiting, when it comes to edge AI (or any technology for that matter), is a mistake. The technology is evolving at a previously unseen, exponential rate, so past tech adoption isn't a good predictor. And disruptors have and will continue to pop up everywhere, shredding business strategies that took years to build and execute, being the Netflix to your Blockbuster.

While it's true that the costs associated with AI can be high, and that ROI will almost certainly take several iterations, you don't have to start big. Start small, but start now.

LOOK AT THE EDGE

AI and machine learning (ML) are increasingly allowing computers to predict consumers' needs before they even know it. Edge computing is now enabling AI on the edge, unlocking a whole new world of possibilities.

Just as the computing evolution was pushed from the mainframe to personal computers, cloud computing, acting as today's "mainframe," is now moving to the edge, and so is AI. Unlike mainframes, however, the cloud still remains relevant. In fact, IoT acts as the cloud's smart extensions, enhancing its capabilities.

But what's so revolutionary about edge AI? Before, data gathered from an IoT device had to be pushed to the cloud where it ran through an algorithm. Then, the results were sent back to the device leading to a delayed response. With edge AI, the algorithms reside on the device, but are still continuously refined in the cloud. Having the algorithm on the device decreases the back-and-forth between the device and central infrastructure, increasing convenience for the user.

Does edge AI already exist? Yes! And you are likely experiencing it on a daily basis. Take your iPhone as an example. Apple provided it with an algorithm to recognize people, allowing phones to be unlocked with a glance. It can also recognize objects and group them, so you can easily find related photos.

A more complex example where AI can literally make a difference between life and death is a self-driving car. Sophisticated algorithms process data right there in the car because there's no time, in heavy rush-hour traffic on a highway, to send data to the cloud, process it and send the results back. The car runs in real-time and makes decisions that allow it to drive all on its own.

Cognitive services provided by the big PaaS providers (Google Cloud, AWS, Azure and so on) are the key to training an AI system.

Cloud tools let the machine learn what you already know. Using our car example, a company like Tesla can provide thousands of examples of hazards that need to be avoided, allowing the predictive algorithm to get smart. Once the algorithms are built, AI can move to the edge. As new data is fed into it, the car can now analyze it and make the decision to break, stop, or change course all within a fraction of a second. Should an algorithm fail to recognize an obstacle, the data captured during the incident will be fed into the AI and ML engine. The algorithms will be refined, and updated versions will be distributed to all cars immediately, so that all upgrade their intelligence at once.

AI AND Y-O-U

Every business has a target audience who cares about their experience, so the potential for AI to have an impact in your business is huge. Not tapping into it is a missed business opportunity today, but also a make-or-break choice for the future. Just imagine, consumer apps that are so smart they can predict what the user wants and needs: a banking app that understands where you are, what you're about to buy and suggest a better value next door, or a shopping app that understands your web searches and buying patterns and can recommend a related product when it hits your price point. Creepy or cool, consumers ultimately won't resist the convenience.

Businesses have a tendency to compare themselves to their competitors, not to the best of the best. Remember that book seller? The one that decided not to be a book seller but a computing powerhouse and is now shaking up the entire retail industry? To realize the power of AI at the edge, you must think bigger than your current parameters or even your market.

Is there time to wait for a sign that your competitors are using AI and jump on the bandwagon later? Maybe, but most likely not. Technology is shaking up industries, and if it hasn't reached yours yet, buckle up, because it's coming.

START NOW. START SLOW.

Getting started doesn't have to mean a giant leap forward. Start small by thinking about how to bring AI into your IoT devices, apps and other interfaces including cash registers and billboards, and delight your customers a little at a time.

Focus on the #1 AI-driven action on your edge that your customers may benefit from the most without alienating them if it is incorrect a significant number of times. Your AI experiments should not result in frustration while you learn. Try pushing small batches of algorithms to your devices to see how you can improve the customer experience. Embrace the data feedback loop. Equate iteration with success.

There's no need to invest millions of dollars to upgrade your systems or bring on new staff before you get started. Run AI as an experiment. Start small and see how it translates into ROI. As you begin to reap the benefits of your efforts, increase the investment. Time and again we have seen some of the most valuable business insights derived from surprisingly small data sets.

FEEDBACK TO THE FUTURE

The key to your "first dance" with AI is the feedback loop. Algorithms are built in the central infrastructure, then fed into IoT devices that can analyze the data they collect. New data gathered by the device is then fed back into the central infrastructure to improve the algorithms — or create new ones.

Returning to our example of autonomous cars, it becomes clear how critical these refining feedback loops are. If a car has an accident, like the infamous Tesla tractor bang up, all of Tesla's cars learn from the experience. While people make the same mistakes over and over, the feedback loop means algorithms learn and improve relentlessly.

AI IS HERE. GET MOVING.

Consider all the ways edge AI has snuck into our daily lives. In just one morning, AI can save the day multiple times. While you were having breakfast your phone alerts you that traffic is bad, and you'll need to leave early to make it in time for your meeting. Before you head out your smart fridge tells you to put milk on the shopping list you plan to tackle at the end of the day. You jump in the car and turn on your music app, which predicts the kind of music you'd like to hear in the morning and starts playing. Pleasing. Convenient. Better. That's edge AI.

The puzzle pieces are falling in place and wide-spread business adoption will happen faster than we can imagine. Unlike previous technology revolutions, the month-over-month or year-over-year growth rates have not just stayed constant. They've been increasing at an exponential rate! The effect has been – and will continue to be – stunning! For any company looking to be around in a decade, the time to figure out edge AI is now.

WOLF RUZICKA has been a leader in the technology industry for more than two decades, and joined EastBanc in 2007 as CEO. He is an entrepreneur, technology advocate, and mentor. He serves on numerous boards and groups, including the Microsoft Customer Advisory Board for Microsoft Azure and the World Economic Forum's "Partnering for Cyber Resilience" initiative. Wolf also served as President of APIphany (now part of Microsoft) and held leadership positions with MicroStrategy, Mercedes Benz, Lufthansa, and more.

How AI Will Take Predictive Analytics to the Next Level

BY FRED JACQUET
TECHNOLOGY EVANGELIST

QUICK VIEW

01. Unless you have been hiding under a rock for the last decade, you know that artificial intelligence is not only the next big thing but that it is actually currently happening.

02. AI takes data insight to action, and boosts edge computing and robotic process automation.

03. Make sure not to miss out how AI can take predictive analytics to the next level.

Since February of 2018, scientists from Google's health-tech subsidiary have pioneered innovative ways of creating revolutionary healthcare insights through artificial intelligence prediction algorithms. Based on the back of a patient's eye scan, their system can make predictions against the patient's risk of experiencing a severe cardiac incident.

To achieve this, they trained a machine learning system with medical data including the age, blood pressure, and smoking habits of about 300,000 patients. Technologies including machine learning and advanced algorithms can now help data scientists see health issues ahead of time.

Far from Philip K. Dick's *Minority Report* where humanity's quest was to rule out bad behavior before it happened with guidance from the miraculous precogs, new AI-powered predictive analytics demonstrates a pragmatic purpose for enterprises to execute on.

The evolution of data analysis since the 1980s is quite amazing. In this early age of data intelligence, we were mainly asking, *"What* happened?*"* To answer this question, IT teams were building statistics and more or less interactive reporting. Then, in the 1990s, we started to talk about analysis based on MS Excel. OLAP (online analytical processing) appeared at the same time. The analysis era aimed to answer the question, "*Why* did this happen?" Then, in the 2000s, companies began to specifically focus on, "What's happening *now*?" The advent of dashboards and scorecards built the monitoring generation, which later gave rise to a new question: "What *will* happen?" To answer this, data analysts first used advanced statistics, data mining, and advanced data analytics. Technologies including machine learning, neural networks, and deep learning now help data scientists face the challenges raised by the need for prediction.

WHAT IS PREDICTIVE ANALYTICS?

Advanced analytics solutions consist of a comprehensive combination of elaborated methodologies, technologies, and infrastructures that take analytic processes over and beyond traditional data processing. Its purpose ranges from making predictions to bringing to light actionable insights.

Advanced analytics' most famous application scope is big data analytics, which aims to unveil patterns, subtle correlations, and trends to empower decision-making processes.

At a higher level, when leveraging data mining methods, complex statistical models, and machine learning technologies, advanced analytics allows for making effective data-based decisions and building sentiment analysis or recommendation systems that lead to predictive analytics.

Predictive analytics should consider these four axes: prediction, speed, business, and accessibility.

1. **Prediction**: Predictive analytics goes beyond the standards that allow for producing simple descriptions. In a few words, descriptions let decisions-makers understand the current state of their business, while predictions empower them to implement action plans knowing what may or may not happen next.

2. **Speed**: The strength of prediction capabilities also lies in the ability to come up with actionable outcomes quickly, compared to a business intelligence batch that could require a whole night, and sometimes even days, to see calculations run their course to the end.

3. **Business**: These analytics are business-oriented by their very nature, far from statistical research where you just search for trends within a dataset about past activities.

4. **Accessibility**: The combination of the three previous axes tends to strengthen the need for business accessibility. The link between the prediction solution and its outcomes, and the decision-makers who will put it into action, requires it to be as simple and as natural as possible.

"Information is the oil of the 21st century, and analytics is the combustion engine."

— PETER SONDERGAARD, GARTNER

MARKET ANALYSIS AND IMPORTANCE

Forrester forecasts a 15% compound annual growth rate for the predictive analytics market through 2021. They observed that within a very large inventory of AI-related innovations, including algorithms and solutions, a surprisingly large portion lies in the open-source community.

In their 2018 Magic Quadrant for Data Science and Machine Learning Platforms, Gartner reveals that on top of historic big players, traditional software editors are shifting from classic descriptive and diagnostic analytics to predictive and prescriptive analytics. Far from the mainstream, advanced analytics are no longer the preserve of Google, Apple, Facebook, and Amazon (GAFA).

Studying the worldwide advanced and predictive analytics software market, IDC reports the trends and momentous evolutions within the analytics market. They state that we are reaching a point where companies have a perfect knowledge that, besides the still relevant value of business intelligence tools, they now have to derive the greatest benefit from the value of "forward-looking analytics" — AKA predictive analytics.

"Prediction is very difficult, especially about the future."

— NIELS HENRIK DAVID BOHR (NOBEL PRIZE FOR PHYSICS IN 1922)

THE SCOPE OF PREDICTIVE ANALYTICS

Each and every modern business domain can gain maximum potential from predictive analytics. In all of them, what is key is that they are in possession of sufficiently large, various, detailed, and reliable historical data that are enriched with the latest types of *"intelligent"* data, including those coming from smartphones, connected devices, sensors, logs, etc. All types of companies must be determined and consistent in their desire to leverage their data in optimizing their processes. This applies equally to fraud detection, processes optimization, costs reductions, market trends anticipation, and the discovery and innovation of new business opportunities, for example.

Let's go through four examples.

1. FINTECH AND BANKING

Regarding the use of a credit card, abnormal behaviors can be revealed by establishing models and observing specific patterns of inappropriate use. This is known as fraud detection, and can also apply to other industries. Social media abounds with data that financial companies are exploiting in order to better understand, get closer to, better satisfy, and ultimately retain their customers. It provides them with extremely valuable data that can help them predict the behavior of the market and their customers.

2. OIL AND GAS

This is a domain that takes a huge advantage of Internet of Things (IoT) — more precisely, IIoT (Industrial IoT). For example, General Electric is using data from vibration and thermography systems and also additional extraneous information to extrapolate data, establish correlations, and make predictions in a marine exploration or extraction site. To explore how AI and IoT synergize, refer to AI and IoT: Taking Data Insight to Action.

3. RETAIL

Prediction systems are quite commonly observed in the retail industry to improve engagement and personalization for consumers, as well as delivery processes or stock management. Combined with marketing research and actions, these systems could also aim to anticipate trends or evaluate the adoption rate of a new brand or service.

4. INDUSTRY

This business domain fully takes advantage of its data, even more so from its IIoT data, implementing predictive maintenance. The smart devices, sensors, and intelligent devices they operate produce tons of data and logs that they leverage through models and advanced algorithms. This is critical for them, as it allows them to optimize their production and obviously reduce risks and cost. But the main point is that IIoT devices allow them to better understand how their systems work and can be maintained — whether they are making industrial machineries, airliners, trucks, or wind turbines.

AI TAKES PREDICTIVE ANALYTICS TO THE NEXT LEVEL

Forrester expresses their conviction that companies that want to leverage AI should start with a predictive analytics machine learning system. They state that machine learning is fundamental to artificial intelligence. Let's see how artificial intelligence can enhance predictive analytics.

HEALTHCARE

Machine learning models using echocardiographic data can greatly improve mortality predictions. The amount of data that physicians have at their disposal is so substantial that they don't have material time to extract their maximum potential.

The Geisinger Health data science team in Danville, Pennsylvania worked on data from 170,000 patients for a total of 330,000 Doppler ultrasonography results. Machine learning models and related algorithms let them significant improve their search field. It pointed out that while 500 measurements derived from echocardiography, six echocardiographic, and four clinical variables were the most important for the five-year estimates.

RETAIL

In the worldwide retail market, Alibaba Group is a real phenomenon. On just one day (Nov. 11, 2017), they made more than $25 billion, which represents roughly 26x that day on Amazon Prime. The strongest activity in terms of payment at this time was 256,000 payments per second.

Alibaba is candidly crediting their success to artificial intelligence.

The Alibaba data scienice lab has crafted a "system which uses real-time online data to predict consumer wants, and the models are constantly updated for each individual through AI to take into account purchase history, browsing history, and online activities." They called it "E-Commerce Brain."

I-ERP

Enterprise Resource Planning (ERP) software solutions provide organizations the ability to manage their business. They integrate and centralize their business data, processes, and workflows across different departments in the enterprise to offer 360-degree views of activities. It includes accounting management, financial, manufacturing, production, sales and distribution, human resources, customer relationships, and more.

Combined with AI technologies, i-ERPs (Intelligent ERPs) are designed to track, route, and analyze these processes, and make predictions. They handle repeatable tasks and utilize voice recognition, machine vision, and natural language processing to interact with humans and assist them.

CONCLUSION

Predictive analytics takes full advantage of enormous technological breakthroughs of the data intelligence era including big data analytics, Internet of Things, cloud, and artificial intelligence. Companies are now aware of the richness of their data and intend to derive the maximum benefit from it to run their business.

Artificial intelligence is grouping together several technological innovations including speech recognition, virtual assistants (AKA bots), machine learning, deep learning, machine vision, biometrics, robotic process automation (RPA), text analytics, and natural language processing. There's no doubt that predictive analytics has gotten stronger and is gaining more credibility in the land of digital transformation.

FRED JACQUET has been working in the Information Technology industry for over two decades.
His experience results from great years within IT companies leading the data intelligence market, as well as assuming roles including Solution Architect, CTO, and Evangelist. His preferred playground is data-oriented, i.e. analytics, IoT, and AI.

diving deeper

INTO **ARTIFICIAL INTELLIGENCE**

twitter

 **@aditdeshpande3**

 **@adampaulcoates**

 **@AndrewYNg**

 **@soumithchintala**

 **@petitegeek**

 **@DrAndyPardoe**

 **@drfeifei**

 **@TessFerrandez**

 **@bobehayes**

 **@jeggers**

books

Understanding Machine Learning: From Theory to Algorithms

Get an introduction to machine learning and the algorithmic paradigms it offers, and get a theoretical account of the fundamentals underlying machine learning.

Neural Networks and Deep Learning

Learn about many of the core concepts behind neural networks and deep learning.

Natural Language Processing With Python

Get a highly accessible introduction to the field of natural language processing.

zones

AI dzone.com/ai

The Artificial Intelligence (AI) Zone features all aspects of AI pertaining to Machine Learning, Natural Language Processing, and Cognitive Computing. The AI Zone goes beyond the buzz and provides practical applications of chatbots, deep learning, knowledge engineering, and neural networks.

IoT dzone.com/iot

The Internet of Things (IoT) Zone features all aspects of this multifaceted technology movement. Here you'll find information related to IoT, including Machine to Machine (M2M), real-time data, fog computing, haptics, open distributed computing, and other hot topics. The IoT Zone goes beyond home automation to include wearables, business-oriented technology, and more.

Big Data dzone.com/big-data

The Big Data/Analytics Zone is a prime resource and community for Big Data professionals of all types. We're on top of all the best tips and news for Hadoop, R, and data visualization technologies. Not only that, but we also give you advice from data science experts on how to understand and present that data.

refcardz

Recommendations Using Redis

Unlike other in-memory data stores, Redis can persist your data to a disk and accommodate a wide variety of data structures. In this Refcard, learn to develop a simple recommendation system with Redis, based on user-indicated interests and collaborative filtering.

Machine Learning

Covers machine learning for predictive analytics, explains setting up training and testing data, and offers machine learning model snippets.

Core Python

Python is an interpreted, dynamically typed language. Python uses indentation to create readable, even beautiful code. With Python's vast array of built-in libraries, it can handle many jobs without the need for further libraries, allowing you to write useful code almost immediately.

courses

Machine Learning on Coursera

Get a broad introduction to machine learning, data mining, and statistical pattern recognition.

Practical Deep Learning for Coders

Learn how to apply deep learning best practices to create state-of-the-art models in computer vision, natural language, and recommendation systems.

Learn With Google AI

Get information and exercises to help you develop your skills and advance your machine learning projects.

NAVIGATING A LIBRARY PERFECTLY WITH NLP

Fewer than a third of DZone readers (29%) are currently experimenting with natural language processing (NLP), a field of technology focused on how computers interact with human languages. Rather than working with logical, machine-readable code, developers working with NLP want computers to interact with humans using the language that humans are most comfortable with. The potential for NLP to make support jobs and analytics easier and more insightful can best be seen here, in the library, a hub of near-endless information. Take a look (in this book!) at the most common uses for NLP below:

LIBRARY

We all know search engines by now: we use keywords for what we're searching for, and the engine returns results from a database that match them. NLP can help search engine algorithms parse more complicated requests to deliver more accurate results. Only 28% of survey respondents are working with NLP and search engines.

NLP can help improve insights by parsing text into useful data. Rather than manually bucketing free text responses, NLP programs can analyze this text, find common patterns or keywords, and create insights. This is a very popular application of NLP, with 60% of DZone respondents using it for this purpose.

Translating one language to another by yourself can be time-consuming, and can easily be inaccurate with common tools today. NLP can help by taking a passage of text and applying rules of grammar from another language to translate it correctly. 21% of DZone respondents are using NLP to translate language.

One of the most-often discussed applications of NLP is the chatbot. Chatbots are programs meant to simulate human conversation with a human. With NLP, many are hoping to introduce chatbots that can help customers with complicated requests, using more than just a keyword search that returns some FAQ pages. Chatbots are popular projects with DZone readers, with 61% of respondents using NLP to create chatbots.

Think of the Executive Summary in this Guide, or an abstract for a research paper. These are designed to give readers a high-level look at what the longer, more detailed work is about. NLP has the potential to summarize complicated text into a few sentences, making it easy for anyone to understand the basic premises of a paper. 31% of DZone readers are using NLP for this purpose.

Executive Insights on the Current and Future State of Artificial Intelligence

BY TOM SMITH
RESEARCH ANALYST AT **DZONE**

QUICK VIEW

01. Keys to a successful AI strategy are a well-defined business problem to solve and a sound data management program in place.

02. Companies benefit from AI by automating business processes and operations thus saving significant time and overhead expense.

03. Thanks to AI, organizations are beginning to see real business value from their data and they are able to do so because GPUs have become affordable.

To gather insights on the state of artificial intelligence (AI), and all of its sub segments -- machine learning (ML), natural language processing (NLP), deep learning (DL), robotic process automation (RPA), regression, etc., we talked to 21 executives who are implementing AI in their own organization and helping others understand how AI can help their business. Specifically, we spoke to:

- Assaf Gad, Vice President and Strategic Partnerships, Audioburst
- Tyler Foxworthy, Chief Scientist, DemandJump
- Patric Palm, CEO, Favro
- Sameer Padhye, CEO, FixStream
- Matthew Tillman, CEO, Haven
- Dipti Borkar, V.P. Product Marketing, Kinetica
- Ted Dunning, Chief Application Architect, MapR
- Jeff Aaron, VP Marketing, Mist Systems
- Ebrahim Safavi, Data Scientist, Mist Systems
- Dominic Wellington, Global IT Evangelist at Moogsoft
- Dr. Nils Lenke, Director, Corporate Research, Nuance Communications
- Mark Gamble, Senior Director of Product Marketing, OpenText
- Sri Ramanathan, Group Vice President of Mobile, Oracle
- Sivan Metzger, CEO and Co-founder, ParallelM
- Nisha Talagala, CTO and Co-founder, ParallelM
- Stuart Feffer, Co-founder and CEO, Reality AI
- Sven Denecken, SVP Head of Product Management, SAP S/4 Hana Cloud
- Steve Sloan, Chief Product Officer, SendGrid
- Simon Crosby, CTO, Swim
- Liran Zvibel, CEO and Co-founder, WekaIO
- Daniel DeMillard, A.I. Architect, zvelo

KEY FINDINGS

1. The keys to a successful AI strategy are to have a **well-defined business problem to solve** and to have a **sound data management program** in place to ensure you have the data to solve the business problem. Truly understand what you are trying to do. Think about the use case. Understand the problem you are trying to solve and the benefit of doing so. The proper definition of the problem will help you determine the optimal type of AI you should use to provide the solution or answer the question.

It all starts with data. ML is an exercise in overcoming variation in data. The amount of data needed depends on the amount of variation. Having a huge number of sources is critical for AI implementation. More accurate and properly formatted data leads to more accurate models.

2. The overarching way companies benefit from AI is by **automating business processes and operations**. Automation saves time, improves accuracy, and frees up workers to use their brains to solve higher-level problems. This improves the efficacy of your operations and provides a significant return on investment. Data has a transformative ability on efficiency and revenue by finding value in vast amounts of data that humans can no possibly see on the magnitude of a 99 percent reduction in events and a 10X improvement in IT productivity.

How many smart decisions could be made if you had the ability to use data to make decisions without fatigue or cost? These are ML candidates. You cannot afford for a human to make decisions that save pennies, but a machine can and the impact on the bottom line can be tremendous.

3. The biggest changes in AI in the past year are that companies are beginning to **see real business value from their data and they are able to do so because GPUs have become affordable.**

The past year has seen many breakthroughs in AI, ML, natural language processing (NLP), and deep learning. Companies are seeing real benefits in how AI can drastically help complex systems adapt, learn, and perform

even as the broader dynamics of human communication changes. When done correctly, it can strengthen customer communications, improve security, provide delivery alerts, yield insights into customers, and optimize programmatic advertising and email based on engagement.

There's hardware available at a lower cost and there's data available. You're able to buy GPUs for a few pennies on the cloud, and this enables machines to make decisions by themselves based on new sources of signals and more data. There are new ways to handle vast amounts of disparate data at great speed.

4. The technical solutions being used most frequently in AI are **TensorFlow, Python, R, and Spark**. A data platform is used to bring data in from across the organization and no one is using a single tool. Everyone tries a bunch of tools and then uses the ones that help them fulfill their particular need. TensorFlow has created a big democratization of the ability to use AI.

5. There are a wide number of real-world problems being solved with AI with those in **automotive, financial services, and healthcare** leading the way. The most frequently mentioned applications were security and compliance.

Training models for autonomous vehicles, medical imaging, homeland security, genomic analysis, risk, and fraud detection. A premier autonomous vehicle company is using machine learning to train and refine algorithms for use in self-driving cars. Genomic analysis is being used as an early detection system for bio-terrorism. In another case, image data from pathological samples are being scanned to detect early stages of cancer.

Several clients in financial services are using applications for fraud detection and loan approval helping to institute an AI practice in the organization so people can focus on operational ML as well as compliance and governance. Some providers have helped banking reduce mortgage approval from 45 days to one day. Replacing repetitive tasks with automation is reducing the amount of time expensive of people doing commodity work.

6. The most common issues preventing companies from realizing the benefits of AI are **data, failure to define the problem you are trying to solve, trying to apply AI where it is not the right tool, silos, and insufficient skillsets**.

Not having a handle on data across silos is a huge problem. When your data house isn't in order, you cannot succeed. There is so much data in so many formats it's overwhelming and complex. It becomes a barrier unless you have the right AI platform that can handle the diversity and amount of data.

It's important that an organization approaches AI from the starting point of, "What problem do we need to solve?" rather than, "Let's do something with AI." A lot of people attempt to use AI/ML where it's not needed. To get value from ML, you really need to understand what's going on and the problem you need to solve.

There is a skillset problem: skills and talent are in short supply. Agencies are no better, so we end up with a blind leading the blind situation. We need more effort around executive education. Arm execs with the right questions to ask. AI solutions providers need to be able to help with execution. Subject matter experts understand datasets and the problem, but don't know how to manage and process data.

7. The biggest opportunities for AI are **reduction of operating costs, automation of repeatable processes, and around call centers and improving customer service**. AI is going to become ubiquitous as there is opportunity in every industry. AI is already driving business model transformation. Manufacturing and supply chain are huge cost centers where costs will be able to be reduced significantly.

Automation is big around cars and trucks. Auto manufacturers don't need to worry about selling more cars, but who they sell to will change. Automation is increasing adoption of ERP workloads in the public cloud leading to faster consumption and reconciliation of data. Humans are being empowered to be more effective in their jobs.

Chatbots are revolutionizing customer support. Digital communications have the potential to feel like actual human conversations and as analytics capabilities become more robust, they will be able to deliver valuable insights into consumer preferences, behavior, sentiment, and intent that can be used to deliver personalized experiences throughout the customer lifecycle.

8. The biggest concerns regarding AI today are around **hype, ethics, and security**. There's a fair amount of hype and excitement. Buzzword fatigue is a concern along with over-inflated expectations. Early experiences may slow down the positive and productive progress we need to make AI seamless.

We can build a place with huge inequalities or we can build a place where AI/ML capabilities make the world a better place. That's a choice we have to make by deciding the types of businesses and business practices we want to support. There's a lot of talk around "AI for good," but simply wishing for the best as we develop new technology is not enough. Society needs to rethink how the future will look so we can be prepared for the changes to public safety and the workforce. We need to ensure AI is used to benefit every member of society rather than just the top corporations.

There is an arms race going on in security. Hackers are becoming more sophisticated and finding increasingly clever ways to bypass safeguards. AI is essential to solving that problem. There's a lot of potential for these systems to detect fraudulent or suspicious activity based on the ability to evaluate massive amounts of data in real time.

9. To be proficient on AI projects, developers need to **understand the fundamentals of data science and be proficient in Python**. A developer needs to know the strengths and weaknesses of AI projects, but they do not need to know all of the math and statistics that go into the background. The difference between a developer and a data scientist will grow smaller as there is a democratization of tools and learning making it easier for people to do this work. In the meantime, developers and data scientists need to spend time talking and understanding the problems they are trying to solve and the algorithms that will help the applications solve the problem.

TOM SMITH is a Research Analyst at DZone who excels at gathering insights from analytics—both quantitative and qualitative—to drive business results. His passion is sharing information of value to help people succeed. In his spare time, you can find him either eating at Chipotle or working out at the gym.

Solutions Directory

This directory contains artificial intelligence and machine learning software, platforms, libraries, and frameworks, as well as many other tools to assist your application security. It provides free trial data and product category information gathered from vendor websites and project pages. Solutions are selected for inclusion based on several impartial criteria, including solution maturity, technical innovativeness, relevance, and data availability.

COMPANY	PRODUCT	PRODUCT TYPE	FREE TRIAL	WEBSITE
Accord.NET	Accord.NET	.NET machine learning framework	Open source	accord-framework.net
AirFusion	AirFusion	AI-powered infrastructure monitoring	N/A	airfusion.com
Alteryx	Alteryx Designer	ETL, predictive analytics, spatial analytics, automated workflows, reporting, & visualization	Available by request	alteryx.com/products/alteryx-platform/alteryx-designer
Amazon Web Services	Amazon Machine Learning	Machine learning algorithms-as-a-service, ETL, data visualization, modeling & management APIs, batch & realtime predictive analytics	Free tier available	aws.amazon.com/machine-learning
Anodot	Anodot	Real time analytics & AI-based anomaly detection	Demo available by request	anodot.com/product
Apache Foundation	Joshua	Statistical machine translation toolkit	Open source	incubator.apache.org/projects/joshua.html
Apache Foundation	Lucene	Text search engine library	Open source	lucene.apache.org/core
Apache Foundation	MADlib	Big data machine learning w/SQL	Open source	madlib.apache.org
Apache Foundation	Mahout	Machine learning & data mining on Hadoop	Open source	mahout.apache.org

COMPANY	PRODUCT	PRODUCT TYPE	FREE TRIAL	WEBSITE
Apache Foundation	MXNet	Deep learning library	Open source	mxnet.incubator.apache.org
Apache Foundation	OpenNLP	Machine learning toolkit for natural language processing	Open source	opennlp.apache.org
Apache Foundation	PredictionIO	Machine learning server	Open source	predictionio.apache.org
Apache Foundation	Singa	Machine learning library creation	Open source	singa.incubator.apache.org/en/index.html#
Apache Foundation	Solr	Information retrieval library	Open source	lucene.apache.org/solr
Apache Foundation	Spark MLlib	Machine learning library for Apache Spark	Open source	spark.apache.org/mllib
Apache Foundation	UIMA	Unstructured data processing system	Open source	uima.apache.org
Artificial Solutions	Teneo Platform	NLI platform for chatbots	Demo available	artificial-solutions.com/teneo
BigML	BigML	Predictive analytics server & development platform	Free tier available	bigml.com
Caffe2	Caffe2	Deep learning framework	Open source	caffe2.ai
Chainer	Chainer	Neural network framework	Open source	chainer.org
CLiPS Research Center	Pattern	Python web mining, NLP, machine learning	Open source	clips.uantwerpen.be/pattern
Cloudera	Cloudera Enterprise Data Hub	Predictive analytics, analytic database, & Hadoop distribution	N/A	cloudera.com/products/enterprise-data-hub.html
DataRobot	DataRobot	Machine learning model-building platform	Demo available by request	datarobot.com/product
Dialogflow	Dialogflow	Chatbot development platform	Free tier available	dialogflow.com
EdgeVerge	Infosys Nia	Artificial intelligence collection & analysis platform	Available by request	edgeverve.com/artificial-intelligence/nia/

COMPANY	PRODUCT	PRODUCT TYPE	FREE TRIAL	WEBSITE
EngineRoom.io	ORAC Platform	AI & deep learning platform	N/A	engineroom.io
Google	TensorFlow	Machine learning library	Open source	tensorflow.org
Grakn Labs	GRAKN.AI	Hyper-relational database for AI	Open source	grakn.ai
Grok	Grok	AI-based incident prevention	Demo available by request	grokstream.com
H2O.ai	H2O.ai	Open source prediction engine on Hadoop and Spark	Open source	h2o.ai
Heaton Research	Encog	Machine learning framework	Open source	heatonresearch.com/encog
IBM	Watson	Artificial intelligence development platform	Free tier available	ibm.com/watson
Intel	Intel nGraph	Framework development library	Open source	ai.intel.com/intel-ngraph
Java-ML	Java-ML	Various machine learning algorithms for Java	Open source	java-ml.sourceforge.net
Kaldi	Kaldi	Speech recognition toolkit for C++	Open source	kaldi-asr.org
Kare	Kare	Knowledge automation solution for CX	Demo available by request	karehq.com
Kasisto	KAI	AI platform for chatbots	N/A	kasisto.com/kai
Keras	Keras	Deep learning library for Python	Open source	keras.io
Marvin	Marvin	JavaScript callback AI	Open source	github.com/retrohacker/marvin
MatConvNet	MatConvNet	Convolutional neural networks for MATLAB	Open source	vlfeat.org/matconvnet
Meya.ai	Meya Bot Studio	Web-based IDE for chatbots	14 days	meya.ai
Micro Focus	IDOL	Machine learning, enterprise search, & analytics platform	Available by request	software.microfocus.com/en-us/products/information-data-analytics-idol/overview

COMPANY	PRODUCT	PRODUCT TYPE	FREE TRIAL	WEBSITE
Microsoft	Cortana Intelligence Suite	Predictive analytics & machine learning development platform	Free Azure account available	azure.microsoft.com/en-us/services/machine-learning-studio
Microsoft	CNTK (Cognitive Toolkit)	Deep learning toolkit	Open source	github.com/Microsoft/CNTK
Microsoft	Azure ML Studio	Visual data science workflow app	Free tier available	studio.azureml.net
Microsoft	Distributed Machine Learning Toolkit	Machine learning toolkit	Open source	dmtk.io
mlpack	mlpack	Machine learning library for C++	Open source	mlpack.org
Natural Language Toolkit	NLTK	Natural language processing platform for Python	Open source	nltk.org
Neura	Neura	AI-powered user retention platform	N/A	theneura.com
Neuroph	Neuroph	Neural network framework for Java	Open source	neuroph.sourceforge.net
OpenNN	OpenNN	Neural network library	Open source	opennn.net
Oryx	Oryx 2	Lambda architecture layers for building machine learning apps	Open source	oryx.io
OutSystems	OutSystems	Low-code platform for building enterprise-grade apps	Available by request	outsystems.com/platform
Progress Software	DataRPM	Cognitive predictive maintenance for industrial IoT	N/A	progress.com/datarpm
Rainbird	Rainbird	Cognitive reasoning platform	Demo available by request	rainbird.ai
RainforestQA	RainforestQA Web App Testing	AI-powered web testing platform	Demo available by request	rainforestqa.com/product/web-app-testing
RapidMiner	RapidMiner Studio	Predictive analytics workflow & model builder	30 days	rapidminer.com/products/studio
RapidMiner	RapidMiner Radoop	Predictive analytics on Hadoop & Spark w/R & Python support	30 days	rapidminer.com/products/radoop

COMPANY	PRODUCT	PRODUCT TYPE	FREE TRIAL	WEBSITE
Salesforce	Einstein	CRM automation & predictive analytics	N/A	salesforce.com/products/einstein/overview
Samsung	Veles	Distributed machine learning platform	Open source	github.com/Samsung/veles
Scikit Learn	Scikit Learn	Machine learning libraries for Python	Open source	scikit-learn.org/stable
Shogun	Shogun	Predictive analytics	Open source	shogun-toolbox.org
Skymind	Deeplearning4j	Deep learning software for Java & Scala	Open source	deeplearning4j.org
spaCy	spaCy	Python natural language processing platform	Open source	spacy.io
Stanford University	Stanford CoreNLP	Natural language processing toolkit	Open source	stanfordnlp.github.io/CoreNLP
Tibco	Tibco Spotfire Data Science	Data science, ETL, predictive analytics, execution workflow design & management	Demo available by request	tibco.com/products/datascience
Torch	Torch	Machine learning framework for use w/GPUs	Open source	torch.ch
UMass Amherst	MALLET	Java library for NLP & machine learning	Open source	mallet.cs.umass.edu
University of Montreal	Theano	Deep learning library for Python	Open source	deeplearning.net/software/theano
University of Waikato	Massive Online Analysis	Data stream mining, machine learning	Open source	moa.cms.waikato.ac.nz
University of Waikato	Weka	Machine learning & data mining for Java	Open source	cs.waikato.ac.nz/ml/weka
Unravel	Unravel	Predictive analytics & machine learning performance monitoring	Available by request	unraveldata.com
Wipro	HOLMES	AI development platform	N/A	wipro.com/holmes
Wit.ai	Wit.ai	Natural language interface for apps	Open source	wit.ai

GLOSSARY

ADVANCED ANALYTICS
A comprehensive combination of elaborated methodologies, technologies, and infrastructures that take analytic processes over traditional data processing by bringing to light actionable insights from data.

ALGORITHMS
A set of rules or instructions given to an AI system, neural network, or other machine to help it learn on its own.

ARTIFICIAL INTELLIGENCE
A machine's ability to make decisions and perform tasks that simulate human intelligence and behavior.

ARTIFICIAL NEURAL NETWORK (ANN)
A learning model created to act like a human brain that solves tasks that are too difficult for traditional computer systems to solve.

CHATBOTS
A chat robot (chatbot for short) that is designed to simulate a conversation with human users by communicating through text chats, voice commands, or both; a commonly used interface for computer programs that include AI capabilities.

CLASSIFICATION
A technique that uses algorithms to let machines assign a category to a data point based on training data.

CLUSTERING
A technique that uses algorithms to let machines group data points or items into groups with similar characteristics.

COGNITIVE COMPUTING
A computerized model that mimics the way the human brain thinks. It involves self-learning through data mining, natural language processing, and pattern recognition.

CONVOLUTIONAL NEURAL NETWORK (CNN)
A type of neural network that identifies and makes sense of images.

DATA MINING
The examination of datasets to discover and "mine" patterns from that data that can be of further use.

DATA SCIENCE
A field of study that combines statistics, computer science, and models to analyze sets of structured or unstructured data.

DECISION TREE
A tree- and branch-based model used to map decisions and their possible consequences, similar to a flow chart.

DEEP LEARNING
The ability for machines to autonomously mimic human thought patterns through artificial neural networks composed of cascading layers of information.

ERP (ENTERPRISE RESOURCE PLANNING)
Software solutions that provide organizations the ability to manage their business by integrating and centralizing their business data, processes, and workflows across different departments in the enterprise to offer a 360-degree view of activities.

I-ERP (INTELLIGENT ERPS)
Combined with AI technologies, i-ERPs are designed to track, route, and analyze these processes and make predictions; they handle repeatable tasks and utilize the capacities of voice recognition, machine vision, and NLP to interact with and assist humans.

MACHINE LEARNING
A facet of AI that focuses on algorithms, allowing machines to learn and change without being programmed when exposed to new data.

MACHINE PERCEPTION
The ability for a system to receive and interpret data from the outside world similarly to how humans use their senses; typically done with attached hardware, e.g. sensors.

NATURAL LANGUAGE PROCESSING
The ability for a program to recognize human communication as it is meant to be understood.

PREDICTIVE ANALYTICS
Combines prediction, speed, business, accessibility, ML, DL, and neural networks; helps answer questions like, "What will happen next?"

RECOMMENDATION
A technique that uses algorithms to help machines suggest a choice based on its commonality with historical data.

RECURRENT NEURAL NETWORK (RNN)
A type of neural network that makes sense of sequential information and recognizes patterns, and creates outputs based on those calculations.

REGRESSION
Regression algorithms help machines predict future outcomes or items in a continuous dataset by solving for the pattern of past inputs, e.g. linear regression in statistics.

SUPERVISED LEARNING
A type of machine learning in which output datasets train the machine to generate the desired algorithms like a teacher supervising a student.

SWARM BEHAVIOR
From the perspective of the mathematical modeler, an emergent behavior arising from simple rules that are followed by individuals that does not involve any central coordination.

UNSUPERVISED LEARNING
A type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses; the most common unsupervised learning method is cluster analysis.

INTRODUCING THE

Open Source Zone

Start Contributing to OSS Communities and Discover Practical Use Cases for Open-Source Software

Whether you are transitioning from a closed to an open community or building an OSS project from the ground up, this Zone will help you solve real-world problems with open-source software.

Learn how to make your first OSS contribution, discover best practices for maintaining and securing your community, and explore the nitty-gritty licensing and legal aspects of OSS projects.

COMMITTERS & MAINTAINERS

COMMUNITY GUIDELINES

LICENSES & GOVERNANCE

TECHNICAL DEBT

Visit the Zone

BROUGHT TO YOU IN PARTNERSHIP WITH Flexera