Cognitive Application Development with DataRPM

OpenEdge PdM Integrator Kit

Anoop Premachandran
Director, Software Engineering, Core Product Group
17 November 2017
Agenda

- Cognitive Application Development or Cognitive Computing
- Progress Cognitive Platform - DataRPM
- Demo
- Look inside OpenEdge PdM Integrator Kit & DataRPM
- How can Progress partners make their applications cognitive with PdM Integrator Kit & DataRPM?
- Q & A
Cognitive Application Development or Computing

- Technology platforms that are based on the scientific disciplines of artificial intelligence and signal processing.
- These platforms encompass machine learning, natural language processing, speech recognition and vision (object recognition), human–computer interaction, dialog and narrative generation among other technologies.
- It represents a broader shift in computing, from a programmatic to a probabilistic approach.
Early adopters of applied AI have a unique opportunity to invent new business models, reshape industries, and build the impossible. Put AI to work — right now

O'Reilly AI Conference 2017
Gartner Hype Cycle for Emerging Technologies
Enterprise cognitive system (ECS)

- Makes a new class of complex decision support problems computable, where the business context is ambiguous, multi-faceted and fast-evolving, and what to do in such a situation is usually assessed today by the business user.

- Is designed to synthesize a business context and link it to the desired outcome and recommends evidence-based actions to help the end-user achieve the desired outcome.

- It does so by finding past situations similar to the current situation, and extracting the repeated actions that best influence the desired outcome.
ECS – Key Characteristics

Adaptive
They must learn as information changes, and as goals and requirements evolve.

Interactive
They must interact easily with users so that those users can define their needs comfortably as well as with devices and cloud services.

Iterative & Stateful
Must “remember” previous interactions and make relevant recommendations without significant iterations from the end-user.

Contextual
Must understand, identify, and extract contextual elements such as meaning, time, location, appropriate domain, regulations, task and goal.
ECS from Progress - DataRPM
There are known knowns.
These are things we know that we know.

There are known unknowns.
That is to say, there are things that we know we don't know.

But there are also unknown unknowns.
There are things we don't know we don't know.

Donald Rumsfeld
Only **20%** of asset failures are common and predictable.

A full **80%** of asset failures were seemingly random and not predicted.
ECS for Predictive Maintenance – Why it matters

**Downtime**
- 43%
  - Unplanned Downtime caused by Equipment Failures

**Cost**
- 50%
  - More Repair Costs to fix a Failed Asset than if the Problem was identified Prior to the Failure

**Productivity**
- 5%
  - Manufacturing Production Capacity lost every year due to Unplanned Downtime

**$**
- $22,000
  - Loss-per-minute for Automotive Downtime alone
The Potential for Industry 4.0

Predictive Maintenance will save companies **$630 billion** by 2025
The Modern Maintenance Maturity Model

MANUAL

Failures are predicted AHEAD OF TIME to perform Maintenance at the optimal Time Window. Machine Learning Model MANUALLY built from only SAMPLES of Data using simple Analytics & Visualization Tools to predict for KNOWN Failures ONLY.

AUTOMATED & SELF-LEARNING

Makes Predictive Maintenance truly AUTONOMOUS by using COGNITIVE Machine Learning to AUTOMATICALLY learn the underlying Mechanisms & Operating Conditions of EACH & EVERY Asset, building INDIVIDUAL Models & CONTINUOUSLY updating to DETECT & PREDICT for not just the Known, but more importantly, the UNKNOWN FAILURES.

Maintenance performed AFTER alarms go off. Often TOO LATE to contain damage.

TIME-BASED Maintenance. Works only for expected “normal” WEAR-&-TEAR but not for the majority of Random Failures.

Performing Maintenance AFTER a Failure has already occurred. High Costs & High Risks.
Need to automate predictive maintenance

THE GOAL
Automate Predictive Maintenance for the Industrial IoT

THE PROBLEM
Underlying Manual Processes Do Not Scale for the IIoT

THE APPROACH
Use Meta Learning to automate Machine Learning to solve PdM autonomously & in parallel at mass-scale

THE SOLUTION
A COGNITIVE Predictive Maintenance & Analytics Platform
Progress DataRPM
An automated Cognitive Platform for Predictive Maintenance Solutions for Asset-based Industries that is Self-Learning

Targeted Use-cases
- To predict RANDOM FAILURES on all assets, allowing Corrective Actions to prevent unplanned downtime & unscheduled maintenance
- To predict Process & Component ANOMALIES on all assets that impact yield, quality & efficiency
More Sensors, Compute & Connectivity for far less

TAILWINDS FOR INDUSTRIAL DIGITIZATION

Thanks to innovations from the smartphone revolution and big data, outfitting industrial equipment with sensors and capturing the immense volumes of data has finally become feasible.
Demo
Use Case: ACME Automobiles

- ACME Automobiles is a premium car engine manufacturer that has four manufacturing plants that produce over 1 Million Engines each annually.
- Locations of ACME Automobile's Manufacturing Plants:
  - Germany
  - USA
  - UK

- Personas:
  - VP Manufacturing; Production Manager; Field Service Manager
Use Case: ACME Automobiles

- Each Plant has four assembly lines to produce car engines:
  - Two for 6-cyl gasoline engines
  - One for 4-cyl gasoline engines
  - One for 6-cyl diesel engines
Use Case: ACME Automobiles
Engine Plant Manufacturing steps

Engine Block from mold

Engine Boring Machine

Engine Assembly Line

Crankshaft Assembly Machine

Camshaft Assembly Machine

Spark Plug Assembly Machine

Piston Assembly Machine

To final inspection of Engines
Use Case: Problem & Proposed Solution

Problem

- ACME is only able to detect problems in quality of the Engines post-facto
  - After all the steps in the Manufacturing Assembly line are completed
- This results in a lot of wastage (material, resources) and re-work (time)
- Customer deliveries are severely impacted

Proposed Solution

- The Sensor data produced by all Assets in Manufacturing Plant staring with Engine Boring Machine till the Camshaft Assembly Machine are analysed to build a Prediction model
- Production Managers and Field Engineers will be able to see the Assets at risk on a Dashboard well in advance to take proactive action
- They will be able to see the factors affecting the Asset performance and predicted state of the Assets
- Number of Engines produced will be estimated and predicted shortage of Engines produced and Predicted Dollars at Risk will be displayed in Dashboard along with the most affected Customers
Switch to Demo …
Mobile App for Remote Field Service Manager
Demo Flow

ACME Production Lines (Mock Data) → DataRPM + OpenEdge PdM Integrator Kit → OE ERP Application

1

2

Other systems (simulated)

- Ticketing
- Inventory
- CRM

Other systems (simulated)

Other systems (simulated)
Internals of the demo system
What is Progress OpenEdge PdM Integrator Kit?

Offering tailor-made for Progress partners to embed PdM capabilities built on Progress DataRPM into their products

1. Domain Centric PdM flow
   ships a standard PdM flow with a defined set of Inputs & Outputs

2. Canonical model
   Asset agnostic model to be consumed by business applications/processes

Extensibility

- The flow & canonical model can be extended as per requirements
- Can enrich the data model with more asset information like age, historic records...
  - write more sophisticated rules using Corticon
  - build better UI/dashboards along with other data like maintenance, resource allocation..
- Extensibility in connector can be leveraged to consume the changes to the output model
DataRPM - Machine Learning

Input Data → Data Inference by Machine Learning → Predictions
DataRPM End to End Flow

- **IloT Sensors**
  - Data

- **Enterprise Asset Management Systems**

- **RDBMS Datasources**

- **Hadoop**

---

**Data Management**
- Data Sync
- Data Lake
- Metadata

---

**Machine Learning & Analytics**
- Spark Engine
- Workflow Builder
- Data Science Recipes
- Meta Learning
- Natural Language
- Visualization

---

**Consumption**
- Insights App
- Discovery App
- PdM Apps
- Admin App

© 2017 Progress Software Corporation and/or its subsidiaries or affiliates. All rights reserved.
DataRPM Patented End-to-End Full Tech Stack

Functional & Data Security

Citizen Data Scientists Business Analysts

Data Scientists Data Engineers / BI Architects

Application Integration

REST Consumption

API Layer

Natural Language (NLQA) Search & Discovery Engine

Meta Learning Engine

Recipe Builder Workflow Builder Visualization Library

Data Science Recipes

DataRPM Unified Data Access Layer

REPL Engine (Zeppelin)

DataRPM Analytics Engine

SIGMA MICRO WORKFLOW EXECUTION ENGINE

DATA CONNECTOR & LOADING MANGER TASK ENGINE EVENT LOADER

Spark (ML & MLLIB) Adaptive Indexing Cheetah (Powered by ES)

DataRPM Core IP DataRPM Proprietary Open Source (DataRPM) Open Source (external)

META DATA STORE (MongoDB)

Hadoop Data Lake

SAP HANA, Teradata, Hadoop/Hive, HBase, MongoDB, Oracle, MySQL, PostgreSQL, SQL Server, DB2, Redshift, CSV, Salesforce, Google Spreadsheet...
DataRPM – Digital Twins
DataRPM – Tooling for Machine Learning
DataRPM – Tooling for Machine Learning
DataRPM – Tooling for Machine Learning
DataRPM – Micro Apps
How can Progress partners make their applications cognitive with OpenEdge PdM Integrator Kit?
Partner Development Lifecycle

- Partner App Development
- Partner App Testing
- Customer Adaptation
- Customer Deployment
Partner App Development Process

Define Use-cases

Partner Application Development

ML Application Development
ML Application Development Activities

Data Modelling and (Re)Training
@ Customer’s site with Customer’s data

Sensor Time-series Data

Data Prep, Data Cleanse, Feature Engg

Unsupervised Learning, Anomaly Detection, Labeling, Meta-Data*

Influencing Factors Analysis – extract human-understandable meta-data*

Domain Expert

Data Scientist

Scoring/Testing Models (Predictions)

Test Sensor Time-series Data

Asset ID | Time      | Sensor | Value
---------|-----------|--------|---------
110001   | 112310901 | 121    | 101.10  
110001   | 112310921 | 122    | 121.10  
110001   | 112310931 | 123    | 113.10  
110001   | 112311901 | 121    | 011.10  
110001   | 112311931 | 121    | 011.10  

Data Prep

Data Cleanse

Feature Engg

Predictive Model

Stored Predictions

Data Scientist

Domain Expert

Train and create predictive model

Re-train on a scheduled basis – 1 month?
Dev Phase activities Summary

Partner Business Analyst
- Comes up with the PdM use-cases

Partner Domain Expert
- Identifies the business actions per asset type.
- Eg: Call customer, raise ticket, escalate via email etc

Partner Application Developer
- Consumes Progress PdM APIs
- Builds business processes against the business actions identified
- Builds custom UI joining predictions with asset information

Partner Data Scientist
- Orchestrates the model creation flow for the PdM usecases
- Leverages the extra information in the partner app (based on schema) and writes custom recipes
- Creates the Scoring flow
Partner Testing Process

1. Deploy integrated application
2. Create Sample model
3. End to End Testing
Customer Adaptation and Deployment

- Build models using actual sensor data
- Map model output with business actions
- Setup scoring flow
- Customer acceptance test
Q & A
Current and Future of Predictive Maintenance

Macro + Micro Patterns approach can only detect and predict for 20% of the failures

“20% Known Unknowns”

Micro Anti-Patterns approach can detect and predict for 80% of the failures

“80% Unknown Unknowns”

Generalized Models created using Samples simply don’t work

Only 20% of Asset Failures are Common & Predictable

80% of Asset Failures are seemingly Random

Samples of Sensor Data from Samples of Assets Generalizes Sample Models to the entire Asset Population

But leads to Poor Results once finally live in Production with ALL assets & diverse Data

Need Individual Predictive Models for each & every Asset to get at that Unknown 80% Millions of Assets = Millions of Models