Predicting Service Outage Using Machine Learning Techniques

HPE Innovation Center
HPE Innovation Center - Our AI Expertise

- Sense
  - Computer Vision
  - Audio/ Speech Processing

- Learn
  - Machine Learning

- Comprehend
  - Natural Language Processing
  - Knowledge Representation

- Act
  - Expert System
IT Operations Analytics (ITOA)
Optimize IT operational service in near real time for production application and infrastructure computing environments.
IT operations analytics (ITOA)

Analyze
high volume structured, unstructured log data and business data

Proactively **avoid** service interruptions, slowdowns and outages
Have **faster** root-cause analysis and problem recovery times
**Enhance** system and application performance
**Improve** end-user experience
**Increase** operational efficiency
**Improve** computing resource utilization
Worldwide IT Operations Analytics Market, 2016

$1.9 Billion

32.9%

Hewlett Packard Enterprise

+17.0% y/y

$97.8M

Note: 2016 Growth (%), and Revenue ($M)
Source: IDC, 2017
Worldwide IT Operations Analytics Business Revenue Share

by region

2016

Americas 74.5%
EMEA 17.5%
APJ 8.0%

Source: IDC, 2017
Worldwide IT Operations Analytics Business Revenue Share

- **Public cloud services**: 13.1%
- **On-premise/others**: 86.9%

**Source**: IDC, 2017

**Deployment type**

**2016**
Service Outage Prediction: a case study
25 TB log data
Probability of occurrence of service outage
Service Outage Prediction

Data Source
- Sys Log
- DB Log
- MW Log
- App Log

Data Pre-processing
- Data cleaning

Log Analysis

Anomaly Detection

Prediction Model

Probability of Outage Occurrence

0 - 1
Service Outage Prediction

Data Source
- Sys Log
- DB Log
- MW Log
- App Log

Data Pre-processing
Data cleaning

Log Analysis

Anomaly Detection of Time Series Data

Numeric Metrics
e.g. DB log

Unstructured Text log
e.g. Sys/MW log

Text Analysis of Logs

Prediction Model
Service Outage Prediction - Three Models

- Naïve Bayes
- Logistic Regression
- Deep Learning
Performance Comparison of Different Models

- CNN
- Logistic Regression
- Naïve Bayes
- Baseline benchmark
Building Deep Learning Model with Imbalanced Data

- Dimension Reduction (PCA)
- Augmentation (Autoencoders)
- Over-sampling (SMOTE)
- Deep Learning (CNN)
- Inference

Input vector

Data

Encoder

Decoder

Synthetic data

Compressed

+ +

- -
Ensembles of CNN

- Dimension Reduction (PCA)
- Augmentation (Autoencoders)
- Over-sampling (SMOTE)
- Deep Learning (CNN)
- Inference
More accurate prediction of the service outage

Higher customer satisfaction

Higher market penetration rate
HPE Portfolio for Deep Learning
HPE has a comprehensive, purpose-built portfolio for Deep Learning

HPE has a comprehensive, purpose-built portfolio for Deep Learning. This portfolio is designed to support the needs of various industries, including government, academia, financial services, life sciences, health, autonomous vehicles/Manufacturing, and more. It includes a range of compute solutions ideal for training models in the data center, compute for both training models and inference at edge, and edge analytics and inference engine.

### Compute ideal for training models in data center

- **HPE SGI 8600** - Petaflop scale for deep learning and HPC
- **HPE Apollo 6500** - The enterprise bridge to accelerated computing
- **HPE Apollo sx40** - Maximize GPU capacity and performance with lower TCO
- **HPE Apollo 2000** - The bridge to enterprise scale-out architecture

### Compute for both training models and inference at edge

- **HPE Edgeline EL4000** - Unprecedented deep edge compute and high capacity storage; open standards

### Edge analytics and inference engine

- **HPE Apollo 4520**
- **HPC Data Management Framework Software** - Large-scale, storage virtualization & tiered data management platform

### AI Software Framework

**Easy Setup and Flexible OS**
Using Bright Computing’s distribution of deep learning software development components and workload management tool integration

### HPC Storage

- **HPE Apollo sx40** - Maximize GPU capacity and performance with lower TCO

### Choice of Fabrics

- **Arista Networking**
- Intel® Omni-Path Architecture
- Mellanox InfiniBand
- HPE FlexFabric Network

---

HPE Confidential
HPE AI customers gaining competitive advantage

Solving complex AI challenges with hybrid cluster

“Through our partnership with SGI, and now HPE, the Tokyo Institute of Technology has worked successfully to deliver a converged world-leading HPC and Deep Learning platform….”

Satoshi Matsuoka
Professor and TSUBAME Leader
Tokyo Institute of Technology..

Upping the ante on Artificial Intelligence

“The discoveries and insights our researchers are now uncovering will have direct effects on human lives by way of advancing precision medicine, increasing energy efficiency, and improving policymaking for the economy.”

Dr. Nick Nystrom
Senior Director of Research
Pittsburgh Supercomputing Center

Libratus AI program defeats world’s best poker players

“The best AI’s ability to do strategic reasoning with imperfect information has now surpassed that of the best humans.”

Dr. Tuomas Sandholm
Professor of Computer Science Department
Carnegie Mellon University

HPE Confidential
HPE demystifies deep learning for faster intelligence

**Get started rapidly:**
Develop deep learning models

**Scale and Integrate:**
Deliver attractive returns

**Optimize Environment:**
Enhance competitive advantage

---

**AI expertise and solutions**
to “get started” with deep learning models

**Proven blueprints and services**
for “scalable” production deployments

**Technology integration**
capabilities to maximize performance

**Expertise**
- Rapid technology selection guides
- State of the art training

**Solutions**
- Integrated purpose-built solutions
- Out of the box solutions

**Proven Blueprints**
- Reference Architectures
- Innovation labs for best practices

**Services**
- Deploy, integrate and support
- Flexible, on-demand capacity

**Integration capabilities**
- Enhanced global Centers of Excellence
- Next gen technology integration

---

New AI expertise, blueprints and technologies to get started, scale, integrate and optimize
AI expertise and solutions to “get started” with deep learning models

Select ideal technologies & systems

Make Informed technology decisions with New HPE Deep Learning Cookbook

Train your teams

Gain organizational competencies with Enhanced Deep Learning Institute

Enhance employee productivity

Accelerate app development with New deep learning integrated solution

Leverage “out of the box” solutions

Increase security of e-commerce with Enhanced HPE Fraud Detection solution

Comprehensive technology selection tool
- Estimates & refines performance
- Characterizes frameworks
- Recommends ideal hardware and software stacks

State of the art deep learning training
- Latest techniques
- Software frameworks
- Infrastructure requirements
- Hands on, instructor led

Pre-configured, proven hardware & software solution
- Purpose-build platform
- Easy to use and install
- Simple management
- Automated framework updates

HPE Fraud Detection Solution with Kinetica
- Uses deep learning techniques
- Qualified with Kinetica in-memory GPU database
- NVIDIA GPU accelerators

Select ideal technologies & systems

New foundation to “get started” with deep learning models

IT Expertise

Solutions
Implement and integrate your production environment

Powering artificial intelligence research

New AI Innovation Labs

Established AI testbed combining both HPE hardware and state of the art AI software to:
- Collaborate with leading academia on AI research projects
- Support internal HPE AI research
- Support select customers and partners in research and POC

Over 75 AI experts in global AI Innovation Labs

Key benefits
- **Stable AI environment** to accelerate time-to-value
- **Purpose-built platforms** to handle the most extreme performance needs
- **Early access** to latest technologies and new, cutting edge hardware and software, often pre-production
- **Faster provisioning and setup**

Advisory services for app and data integration

Expertise for quick deployments
- Implement and integrate with a trusted partner, HPE Pointnext services
- Leverage advisory, professional and operational services
HPE Early Access program for 8-way SXM2 GPU Volta!
Accelerating Deep Learning adoption

First Tier One OEM to provide early access to 8-way Volta SXM2 server

Timing is Q12018 for select customers

Configuration, benchmarking and technology selection guidance

Limited engagement for select customers through your Sales contacts

Advice, planning, design, benchmarking
Thank You