Log-based Predictive Maintenance

Ruben Sipos, Cornel U.
Dmitriy Fradkin, Siemens
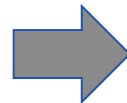Fabian Moerchen, Amazon
Zhuang (John) Wang, **Skytree**

(work performed while all authors were employed by Siemens Research in Princeton, NJ)

Log-based Predictive Maintenance

• Helps determine the running condition of in-service equipment to predict when and where repairs should be performed
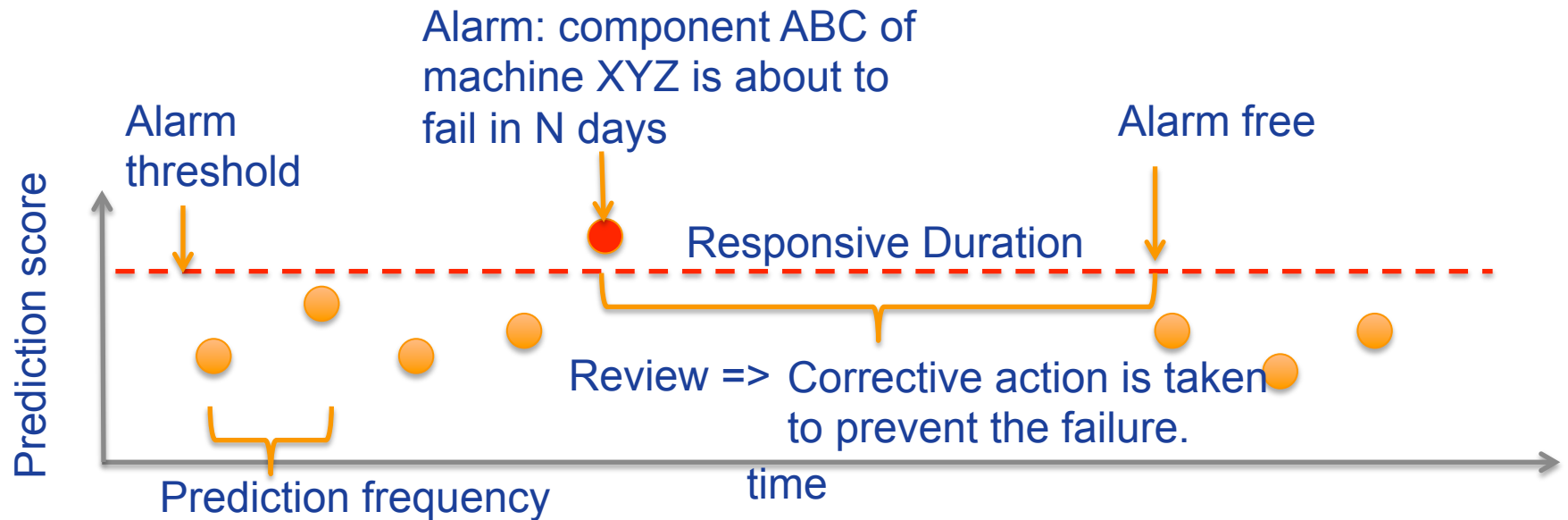

Logs from various software applications record detailed operations of equipment

```
2012-09-08  21:21:09    CT_ITH  2    IRS scan information: Topo   tot Rdg: 6460, defects: HV-Drop 0, HT
2012-09-08  21:21:07    CT_IS   320  First data from IRS arrived in IS
2012-09-08  21:21:06    CT_CHR  107  SKIP button is hidden (bolus exam not startedactive entry is not m
2012-09-08  21:21:06    CT_CHR  107  CANCEL button is dimmed (scanstate == UiScanning)
2012-09-08  21:21:06    CT_CHR  107  CANCELMOVE button is hidden (scanstate == UiScanning)
2012-09-08  21:21:06    CT_STC  93   Zero button has been locked.
2012-09-08  21:21:06    CT_ISV  49   Timer was started waiting for .......... some answer from IRS Recc
2012-09-08  21:21:06    CT_CHR  107  Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT REGISTRATION EXAM#
2012-09-08  21:21:06    CT_ISV  49   Timer was started waiting for .......... SCAN_DONE message for 778
2012-09-08  21:21:05    CT_ACU  3037    Control info ACU   (E c0 03 25 d4 4a 00 00)
2012-09-08  21:21:03    CT_ACU  3037    Control info ACU   (E c0 03 25 d3 7e 00 00)
2012-09-08  21:21:02    CT_TG   9    A new planning image of type CLEAR SEGMENT with series loid EMPTY
2012-09-08  21:21:02    CT_MSM  211  (+) Receiving MeasStart request (scan 0 of range 0)
2012-09-08  21:21:02    CT_SSQ  493  received IS-Notification: SsqMeasISClient::onIrsStarted / IS_MsgId
2012-09-08  21:21:02    CT_SIS  120  &Start button released&
2012-09-08  21:21:02    CT_SIS  85   Multiplexer: the START-button on control box has been released.
2012-09-08  21:21:02    CT_ISV  46   Receiving IsvMsgStartAck message of size 80 Bytes from IRS.
2012-09-08  21:21:02    CT_CHR  107  SKIP button is hidden (scanstate == UiStartedTopobolus exam not st
2012-09-08  21:21:02    CT_CHR  107  CANCEL button is dimmed (scanstate == UiStartedTopo)
2012-09-08  21:21:02    CT_CHR  107  CANCELMOVE button is hidden (scanstate == UiStartedTopo)
2012-09-08  21:21:02    CT_CHR  107  LOAD button is hidden (suspend-button is shown)
2012-09-08  21:21:02    CT_CHR  107  SUSPEND button is ENABLED
2012-09-08  21:21:02    CT_CHR  107  Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT REGISTRATION EXAM#
2012-09-08  21:21:02    CT_ISV  45   Sending IsvMsgStart message of size 72 Bytes to IRS.
```

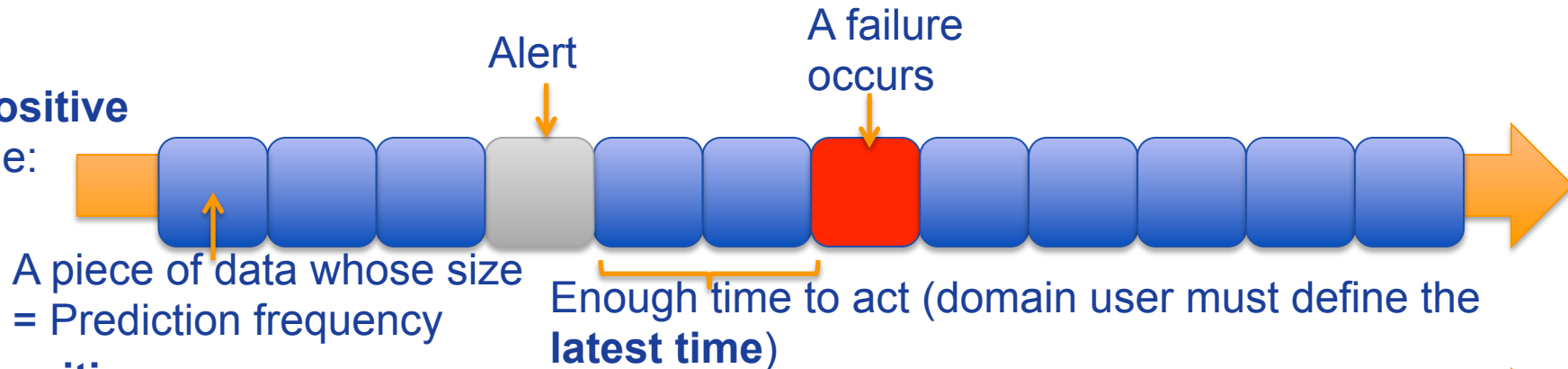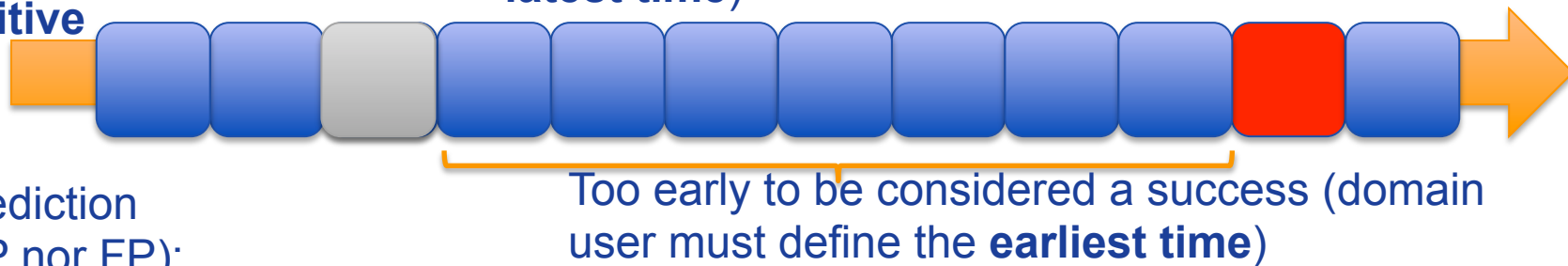Mining logs helps in detecting potential issues in advance.

Maintenance Service Process and Data Flow

How is Model Used in Real-life?

How is Model Evaluated in Real-life?

True positive example:

Alert

A failure occurs

A piece of data whose size = Prediction frequency

Enough time to act (domain user must define the **latest time**)

False positive example:

Too early to be considered a success (domain user must define the **earliest time**)

Invalid prediction (neither TP nor FP):

Too late to be considered a success; but algorithm did capture a failure

Alerts should be muted in **Responsive Duration**

How is Model Evaluated in Real-life? (cont.)

Counting of TP/FP:

3 true alerts but only count 1 TP (1 failure can only have at most 1 TP)

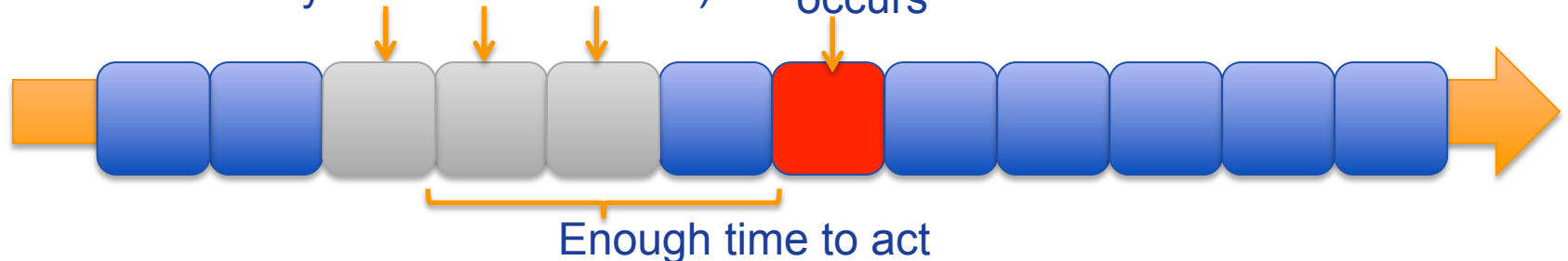A failure occurs

Enough time to act

2 false alerts but only count 1 FP (1 Responsive Duration can only have at most 1 FP)

Too early to be treated success

- **Recall:** TP/ All failures
- **Precision:** TP/ (TP+ FP)
- **Predictive-Maintenance-based AUC** (**PM-AUC**): computed like regular PR-AUC but using the above defined TP and FP.

Application Event Logs: the X

- Logs contain every detail of how an equipment was being operated.
- Logs contain complicated data types.
- We use bag of "word" representation to featurize logs, where "word" is captured by diff. feature extractor.

Time stamp **Eventcode** **Message text**

event

Temporal sequence

Categorical variables

Encoded Time series

Unstructured text

```
2012-09-08  21:21:09   CT_ITH   2    IRS scan information: Topo    tot Rdg: 6460, defects: HV-Drop 0, HT
2012-09-08  21:21:07   CT_IS    320  First data from IRS arrived in IS
2012-09-08  21:21:06   CT_CHR   107  SKIP button is hidden (bolus exam not startedactive entry is not n
2012-09-08  21:21:06   CT_CHR   107  CANCEL button is dimmed (scanstate == UiScanning)
2012-09-08  21:21:06   CT_CHR   107  CANCELMOVE button is hidden (scanstate == UiScanning)
2012-09-08  21:21:06   CT_STC   93   Zero button has been locked.
2012-09-08  21:21:06   CT_ISV   49   Timer was started waiting for .......... some a       IRS Rec
2012-09-08  21:21:06   CT_CHR   107  Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT        EXAM#
2012-09-08  21:21:06   CT_ISV   49   Timer was started waiting for .......... SCAN_          or 778
2012-09-0        ACU   3037     Control info ACU    (E c0 03 25 d4 4a 00 00
2012-09-0        ACU   3037     Control info ACU    (E c0 03 25 d3 7e 00 00)
2012-09-0        TG    9    A new planning image of type CLEAR SEGMENT with series loid EMPTY
2012-09-08  21:21:02   CT_MSM   211  (+) Receiving MeasStart request (scan 0 of range 0)
2012-09-08  21:21:02   CT_SSQ   493  received IS-Notification: SsqMeasISClient::onIrsStarted / IS_MsgIc
2012-09-08  21:21:02   CT_SIS   120  &Start button released&
2012-09-08  21:21:02   CT_SIS   85   Multiplexer: the START-button on control box has been released.
2012-09-08  21:21:02   CT_ISV   46   Receiving IsvMsgStartAck message of size 80 Bytes from IRS.
2012-09-08  21:21:02   CT            is hidden (scanstate == UiStartedTopobolus exam not st
2012-09-08  21:21:02   CT            on is dimmed (scanstate == UiStartedTopo)
2012-09-08  21:21:02   CT            button is hidden (scanstate == UiStartedTopo)
2012-09-08  21:21:02   CT_CHR   107  LOA   ton is hidden (suspend-button is shown)
2012-09-08  21:21:02   CT_CHR   107  SUSP  D button is ENABLED
2012-09-08  21:21:02   CT_CHR   107  Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT REGISTRATION EXAM#
2012-09-08  21:21:02   CT_ISV   45   Sending IsvMsgStart message of size 72 Bytes to IRS.
```

Service Data: the Y

- Service data records details of performed services:
    - service open/close date
    - equipment id
    - component replacement
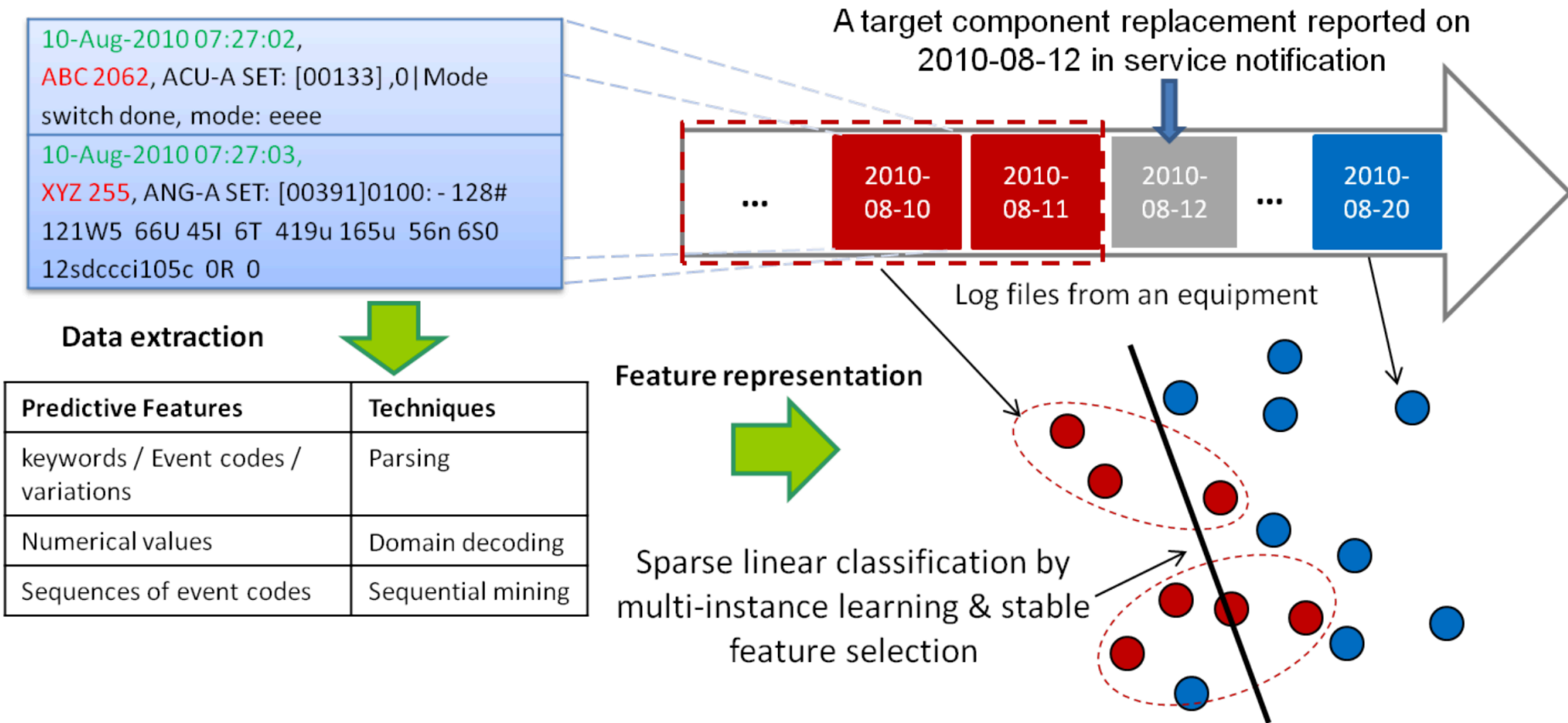   - …

23	Equipment	Notification Number	Notification date	Notification time	Text short	Ausfalldatum ab	Ausfallzeit ab	Ausfalldatum bis	Ausfallzeit bis	Division IPB		Material		IBCKF_M	IBCKF_MVKZ	
24	Gesamtergebnis													5,263	5,071,584.67	3,301
25	14003	728100456470	07.09.2007	12:41:33	GENERATOR E	07.09.2007	12:41:33	07.09.2007	23:00:00	01	AXA	05764506	MEGALIX Cat 125	1	17,778.75	
26		728100543184	21.02.2008	15:38:47	SOFTWARE A	21.02.2008	15:38:47	06.03.2008	22:38:34	01	AXA	04776063	HR Heliflex Optik o	1	2,698.06	
27												05896852	CCR Board D71	1	1,547.29	
28												07149979	Kamera Head TH	1	5,081.49	
29		728100546276	27.02.2008	11:26:07	SOFTWARE H	27.02.2008	11:26:07	22.03.2008	02:56:31	01	AXA	07128866	CPU_D10_IS_OPE	1	1,077.50	
30		728100963960	15.12.2009	11:22:05	X-RAY NO X-R	15.12.2009	11:22:05	15.12.2009	20:00:00	01	AXA	05997817	PCA QUAD DSP1	1	1,077.50	
31	14004	400101791908	19.01.2007	16:38:12	CSE/27068 RE	19.01.2007	16:38:12	25.01.2007	20:24:00	01	AXA	01192124	ABSTANDSHALTE	2	5.52	
32		400101828175	28.02.2007	11:50:16	NO XRAY;RO	28.02.2007	11:50:16	29.06.2007	16:30:00	01	AXA	04775990	D90 Ein/Aus (Logi	1	534.44	
33												05246264	CBL Signl Comm D	4	206.88	
34												05998377	ASM Host 3 Modu	1	2,432.56	
35												07325579	ASM IMPAC Maste	1	2,870.46	
36												07325900	ASM IPS2 WITH M	1	4,999.60	
37												07326080	PCA Power ON 2	1	724.08	
38		400101917799	05.06.2007	10:00:09	B-PLANE CAR	05.06.2007	10:00:10	09.06.2007	13:10:00	01	AXA	06465590	Potentiometer 354	2	108.62	
39												10140940	Schaltleiste Bx/dB	2	698.22	
40		400101950664	11.07.2007	17:56:41	CD WRITER DC	11.07.2007	17:56:41	26.07.2007	16:40:00	01	AXA	01768535	WANDLER;AC-DC	1	208.17	
41												10051975	Adapter IDE to US	2	56.90	
42												10051978	Drive DVD/CD R/W	1	140.94	
43		400102027866	03.10.2007	12:24:11	A-PLANE FLU	03.10.2007	12:24:11	04.10.2007	09:40:00	01	AXA	07721603	FS Mono Artis Tis	1	1,383.51	
44		400102247951	05.03.2008	17:06:54	ECC RAIL CLA	05.03.2008	17:06:54	06.03.2008	11:30:00	01	AXA	04775735	Zubehoer-Schiene	1	76.72	
45		400102260600	13.03.2008	10:33:50	TUBE STARTU	13.03.2008	10:33:51	14.03.2008	10:40:00	01	AXA	07124139	Anlassgeraet N75	1	2,736.85	
46		400102598104	21.10.2008	15:57:59	CSE COLE REF	21.10.2008	15:57:59	28.11.2008	09:10:00	01	AXA	04776063	HR Heliflex Optik o	1	2,775.64	

- **Not always correct but reasonable assumption:**
    a component replacement is the consequence of the component failure

We can use historical service data to correlate any known component failure with its corresponding equipment, time and relevant logs.

The Methodology

10-Aug-2010 07:27:02,
ABC 2062, ACU-A SET: [00133] ,0|Mode
switch done, mode: eeee

10-Aug-2010 07:27:03,
XYZ 255, ANG-A SET: [00391]0100: - 128#
121W5 66U 45I 6T 419u 165u 56n 6S0
12sdccci105c 0R 0

Data extraction

Predictive Features	Techniques
keywords / Event codes / variations	Parsing
Numerical values	Domain decoding
Sequences of event codes	Sequential mining

A target component replacement reported on 2010-08-12 in service notification

... | 2010-08-10 | 2010-08-11 | 2010-08-12 | ... | 2010-08-20

Log files from an equipment

Feature representation

Sparse linear classification by multi-instance learning & stable feature selection

We propose a simple but effective algorithm for the resulting MIL problem with imbalanced labels and high-dimensional features

Domain-based Evaluation Results

- Data

	Dataset C	Dataset A
# instances	181,945	422,560
# failures	88	108
# features	6,664	14,367

- Competing MIL algorithms
  - AllInstance (ICML, 2005)
  - Aggregated
  - MILES (PAMI, 2006)
  - MI-SVM (NIPS, 2003)

PM-AUC comparison (bag-level 5 CV with stratified sampling):

	C	A
random	0.037 (0.004)	0.017 (0.003)
AllInst.	0.293 (0.014)	0.620 (0.013)
Agg.	0.174 (0.013)	0.498 (0.016)
MILES	0.170 (0.011)	0.427 (0.117)
MI-SVM	0.216 (0.038)	0.700 (0.014)
Ours	**0.319 (0.015)**	**0.730 (0.011)**

PM-ROC with Different Model Complexity

70% precision 25% recall with 7-day predictive interval

Simple model can be reviewed for root cause analysis

70% precision 80% recall with 7-day predictive interval

AUC:0.210, #features:9
AUC:0.251, #features:176
AUC:0.319, #features:379

AUC:0.368, #features:9
AUC:0.691, #features:297
AUC:0.730, #features:389
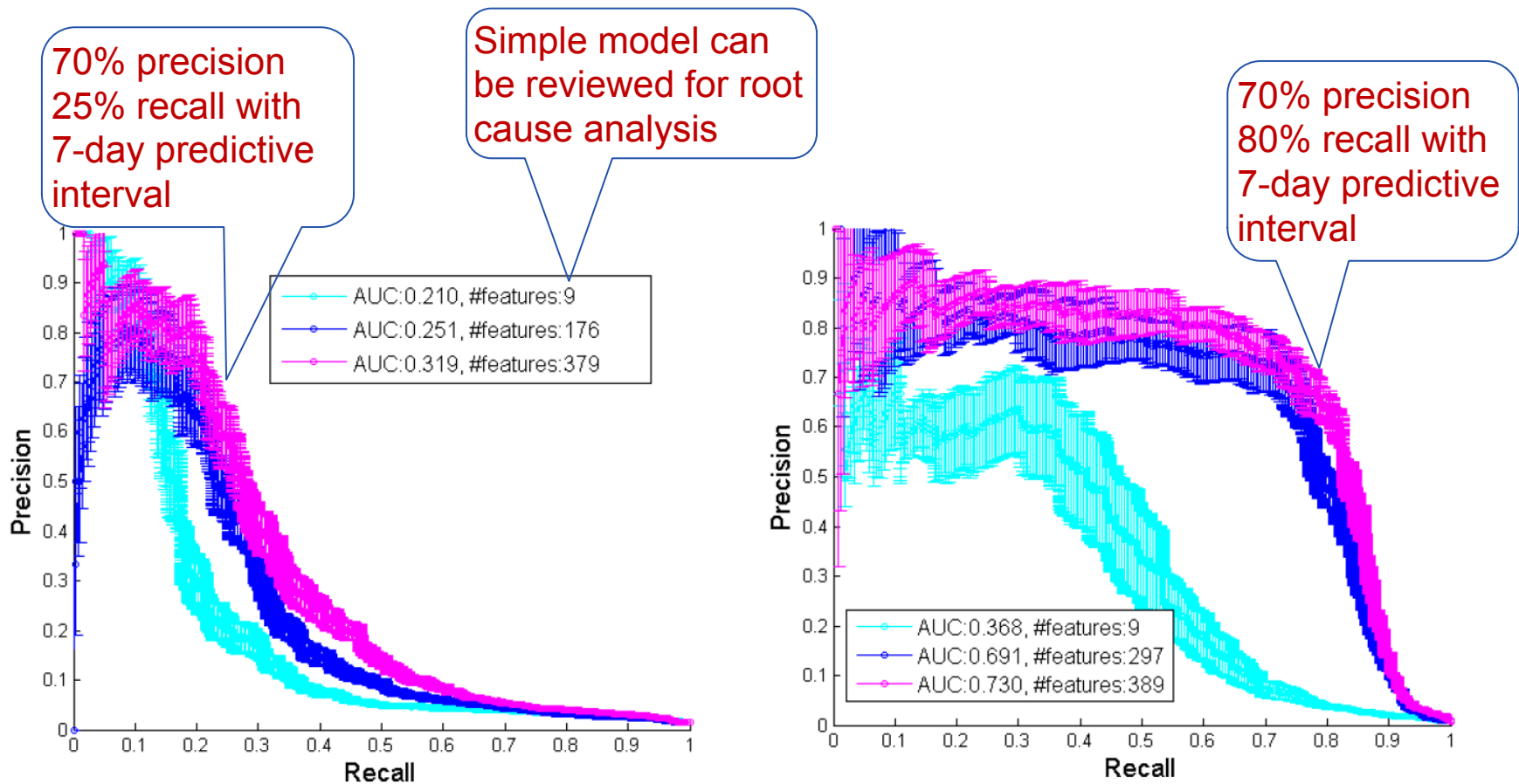
See more results in paper:
- **On Utility of Single Features**
- **Local vs Global Models**

Deployment

- Guided the client in implementing the approach on SAS platform.

Re-implemented the approach in SAS EM workflow
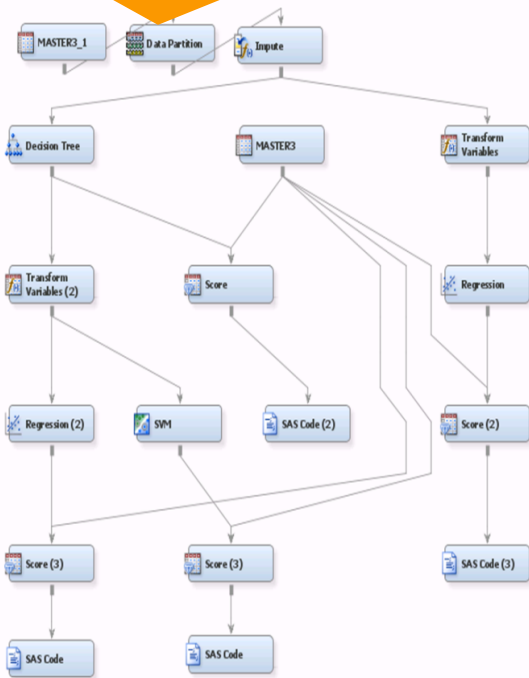
User portal for decision support

alert scores
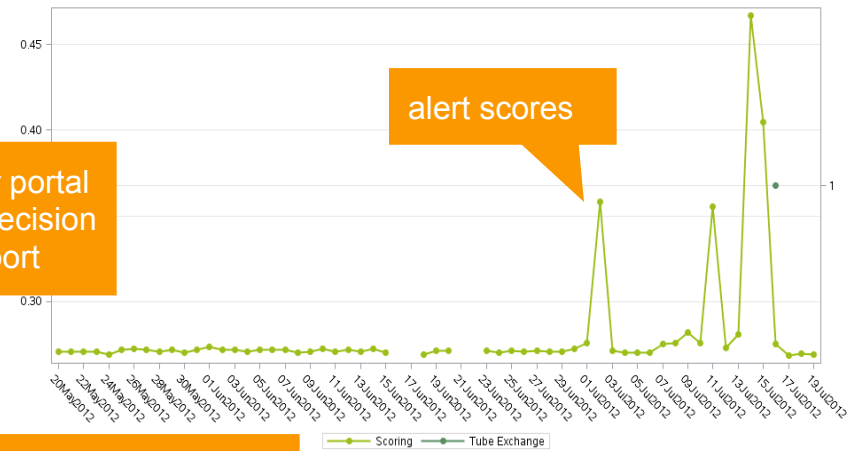
Predicted alerts for all the monitored equipments stored in a central DB

Thank you for your time! Questions?

End-to-end Predictive Maintenance App available at *Skytree, Inc*