



Universidade Federal do Rio de Janeiro

Escola Politécnica

MBA em Big Data, Business Intelligence e Business Analytics  
(MB3B)

**BUSINESS INTELLIGENCE NA MANUTENÇÃO PREDITIVA**

Autor:

---

Rayane de Sousa, B. EM.

Orientador:

---

Manoel Villas Boas Junior, M. Sc.

Coorientador:

---

Edilberto Strauss, Ph. D.

Examinador:

---

José Airton Chaves Cavalcante Junior, D. Sc.

Examinador:

---

Vinicius Drumond Gonzaga, M. Sc.

Examinador:

---

Flávio Luis de Mello, D. Sc.

**Rio de Janeiro**  
**Novembro 2021**

## Declaração de Autoria e de Direitos

Eu, **Rayane de Sousa** CPF 145.788.527-17, autor da monografia **BUSINESS INTELLIGENCE NA MANUTENÇÃO PREDITIVA**, subscrevo para os devidos fins, as seguintes informações:

1. O autor declara que o trabalho apresentado na defesa da monografia do curso de Pós-Graduação, Especialização MBA em Big Data, Business Intelligence e Business Analytics da Escola Politécnica da UFRJ é de sua autoria, sendo original em forma e conteúdo.
2. Excetuam-se do item 1 eventuais transcrições de texto, figuras, tabelas, conceitos e idéias, que identifiquem claramente a fonte original, explicitando as autorizações obtidas dos respectivos proprietários, quando necessárias.
3. O autor permite que a UFRJ, por um prazo indeterminado, efetue em qualquer mídia de divulgação, a publicação do trabalho acadêmico em sua totalidade, ou em parte. Essa autorização não envolve ônus de qualquer natureza à UFRJ, ou aos seus representantes.
4. O autor declara, ainda, ter a capacidade jurídica para a prática do presente ato, assim como ter conhecimento do teor da presente Declaração, estando ciente das sanções e punições legais, no que tange a cópia parcial, ou total, de obra intelectual, o que se configura como violação do direito autoral previsto no Código Penal Brasileiro no art.184 e art.299, bem como na Lei 9.610.
5. O autor é o único responsável pelo conteúdo apresentado nos trabalhos acadêmicos publicados, não cabendo à UFRJ, aos seus representantes, ou ao(s) orientador(es), qualquer responsabilização/ indenização nesse sentido.
6. Por ser verdade, firmo a presente declaração.

Rio de Janeiro, \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

---

Nome Completo

## **UNIVERSIDADE FEDERAL DO RIO DE JANEIRO**

Av. Athos da Silveira, 149 - Centro de Tecnologia, Bloco H, sala - 212,  
Cidade Universitária Rio de Janeiro – RJ - CEP 21949-900.

Este exemplar é de propriedade Escola Politécnica da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazenar em computador, microfilmear ou adotar qualquer forma de arquivamento.

Permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es).

## **DEDICATÓRIA**

Dedico este trabalho a todos os estudantes e pesquisadores

## **AGRADECIMENTO**

Primeiramente agradeço a minha família, por me incentivar a estar sempre estudando e graças a eles procuro sempre conhecer novos assuntos

Agradeço a meu esposo Vitor, que sempre acreditou em meu potencial e não me deixou desistir sempre estando ao meu lado e também aos meus amigos Luis, Sonia e Mariana sempre acompanhando meu crescimento em meus estudos e me ajudando sempre quando necessário com suas opiniões.

Agradeço ao professor Claudio Miceli pela paciência e dedicação que teve comigo durante toda orientação desse trabalho.

Obrigada a todos.

## RESUMO

Para a indústria 4.0 a coleta de Big Data é muito desafiadora. Serão apresentadas algumas técnicas promissoras de diagnóstico que levam como base o aprendizado de máquinas. O objetivo deste trabalho é aplicar as técnicas de Business Intelligence na Manutenção Preditiva dentro da indústria 4.0. A importância do uso dessas técnicas dentro da indústria, sendo mais preciso, na manutenção preditiva, é a redução de gastos com aperfeiçoamentos, aumento de vida e da performance da máquina. Através de pesquisa exploratória, utilizamos as técnicas de Business Intelligence junto a base de dados que possui informações próximas à realidade dentro da indústria, manifestando assim o desafio do uso dessas técnicas. O resultado obtido durante o processo não foi satisfatório. Isso deixa uma abertura para que o problema persista em trabalhos futuros, os quais sejam realizados utilizando novas técnicas de BI. Concluímos que é possível a utilização dessas técnicas dentro da indústria 4.0 quando direcionada corretamente para cada situação.

Palavras-Chave: Aprendizado de Máquina, Manutenção, Industria 4.0, Curva PF, Inteligencia Artificial.

## **ABSTRACT**

For the 4.0 Industry, gathering Big Data is very challenging. We will present some of the promising diagnosis techniques based on machine learning. The purpose here is to present how to reach expected results when Business Intelligence is applied on the Predictive Maintenance, to solve problems from the 4.0 Industry. The significance of using these techniques within the industry, being more precise in predictive maintenance, is the reduction of expenses with improvements, increase of life and performance of the machine. Through exploratory research, we use Business Intelligence techniques along with a database that has information close to reality within the industry, exhibiting that using these techniques are very challenging. The result obtained during the process was not satisfactory. This leaves an opening for the problem to persist in future work, which is carried out using new BI techniques. In conclusion, it is possible to use these techniques within Industry 4.0 when correctly targeted for each situation.

Keywords: Machine Learning, Maintenance Predictive, Industry 4.0, Curve PF, Business Intelligence.

## SIGLAS

**BI** Business Intelligence

**IA** Inteligência Artificial

**IoT** Internet of Things – Internet das coisas

**ML** Machine Learning

**Np** Numpy

**Plt** Matplotlib.pyplot

**Sns** Seaborn

## LISTA DE FIGURAS

Figura 1	Fases das Revoluções Industriais	2
Figura 2	Diagrama de manutenção	3
Figura 3	Curva PF	11
Figura 4	Curva T x C	11
Figura 5	Manutenções	12
Figura 6	Diagrama de Pareto	13
Figura 7	Instalando as bibliotecas iniciais	16
Figura 8	Lendo o arquivo csv	16
Figura 9	Imprimindo informações do DataFrame	17
Figura 10	Renomeando as colunas do DataFrame	18
Figura 11	Importando outras bibliotecas	18
Figura 12	Correlação existentes na tabela	19
Figura 13	Gráfico de distribuição	19
Figura 14	Gráfico de calor	20
Figura 15	Gráfico random seed	21
Figura 16	Organização dos dados para treino e teste	22
Figura 17	Resultados	23

## LISTA DE TABELAS

Tabela 1	Resultado do DataFrame	17
----------	------------------------	----

# Sumário

<b>Capítulo 1: Introdução.....</b>	<b>1</b>
1.1 – Business Intelligence e Machine Learning na Manutenção Preditiva.....	5
1.1.1 – Business Intelligence.....	5
1.1.2 – Machine Learning.....	6
1.2 – Justificativa.....	8
1.3 – Objetivos.....	8
1.4 – Delimitação.....	8
1.5 – Metodologia.....	9
1.6 – Descrição.....	9
<b>Capítulo 2: Embasamento Teórico.....</b>	<b>10</b>
2.1 – Fundamentos da Indústria 4.0.....	13
<b>Capítulo 3: Experimento e Propostas Tecnológicas.....</b>	<b>14</b>
<b>Capítulo 4: Conclusão e Trabalhos Futuros.....</b>	<b>24</b>
<b>Referências Bibliográficas.....</b>	<b>25</b>
<b>Apêndice 1.....</b>	<b>27</b>
<b>Apêndice 2.....</b>	<b>33</b>

# CAPÍTULO 1

## Introdução

Desde o início da Revolução Industrial, em 1760, houve grandes descobertas sendo assim um período de muitos desenvolvimentos tecnológicos não só na Inglaterra, mas no mundo todo. A descoberta da máquina a vapor garantiu o surgimento da indústria, consolidando a formação do capitalismo, causando grande transformação na economia mundial, estilo de vida, produção de mercadorias, exploração dos recursos naturais e nas relações de trabalho.

Com o passar dos anos, começamos a utilizar muito a tecnologia em nosso dia a dia, trazendo modernização na produção: máquinas mais eficientes; comunicação: utilização de aparelhos eletrônicos como computadores, celulares e câmeras; e nos estudos: o aparecimento da internet, onde hoje é utilizada por quase toda população sendo por computador e/ou por celular. Cada fase da Revolução Industrial têm sua característica, veremos em seguida cada uma delas e suas diferenças:

**Primeira Revolução Industrial:** processo de evolução tecnológica no século XVIII (entre 1760 e 1850). Fase marcada pela máquina a vapor, eólica e hidráulica; substituição da manufatura para maquinofatura e a existência de novas relações de trabalho. Como resultado houve modificação na produção diminuindo o tempo e o aumento dela, as invenções dão possibilidade ao melhor escoamento de matéria prima e também de consumidores favorecendo a distribuição dos bens produzidos.

**Segunda Revolução Industrial:** processo de evolução na metade do século XIX tendo seu fim durante a Segunda Guerra Mundial. Houve avanço da industrialização nos limites geográficos da Europa Ocidental. Essa fase é marcada pelo aperfeiçoamento tecnológico das tecnologias já existentes na primeira fase, vivenciando novas criações aumento ainda mais a produtividade e seus lucros, também houve grande incentivo as pesquisas no ramo da medicina. As principais invenções foram associadas ao uso do petróleo como sua fonte de energia, como o motor a combustão. A eletricidade passou a ser usada em funcionamento de motores, em seu destaque motores elétricos e à explosão, e o aço começou a substituir o ferro.

**Terceira Revolução Industrial:** essa fase também é conhecida como Revolução Tecnocientífica, teve seu início na metade do séc XX. Nessa fase teve crescimento no setor industrial, tecnológico havendo avanço científico espalhando para todo o mundo. As transformações tecnocientíficas são vivenciadas até o dia de hoje na biotecnologia, robótica, avanços na área da genética, telecomunicações, eletrônica, transporte, entre outras áreas, transformando também o modo de vida da sociedade e o espaço geográfico. Esses avanços na área científica também é conhecido como globalização.

**Quarta Revolução Industrial:** é caracterizada pela automação industrial, onde integra-se diferentes tecnologias como inteligência artificial, robótica, internet das coisas, computação em nuvem. Hoje na indústria 4.0 a tecnologia passa por adaptações acompanhando as técnicas preditivas, como: análise de vibrações, ultrassom, inspeção visual, técnicas de análise não destrutivas, entre outros, além de acompanhamentos periódicos.

Figura 1 – Fases das Revoluções Industriais



Fonte: <http://especiais.estadao.com.br/mundodigital/ola-mundo/>

Com isso, antecipamos a necessidade de manutenção, eliminamos o trabalho de desmontagem aumentando o tempo de vida do equipamento, reduzindo as ocorrências de emergência não planejadas impedindo o aumento dos danos. Está centrada na confiabilidade e tem por função assegurar que todos os ativos continuem cumprindo as funções que lhes foram especificados.

Um dos desafios da indústria 4.0 é manter em dia a manutenção preditiva quando se trata de coletar dados das máquinas para o software, sem que haja alguma interferência. Passa que seja possível utilizarmos o *Business Intelligence (BI)* e *Machine Learning (ML)*. O *BI* com suas técnicas para visualizar os dados que foram coletados e também na tomada de decisão, e o *ML* para ensinar a máquina a forma certa dela pensar através de diversas técnicas de aprendizados e usando a melhor técnica para a situação específica.

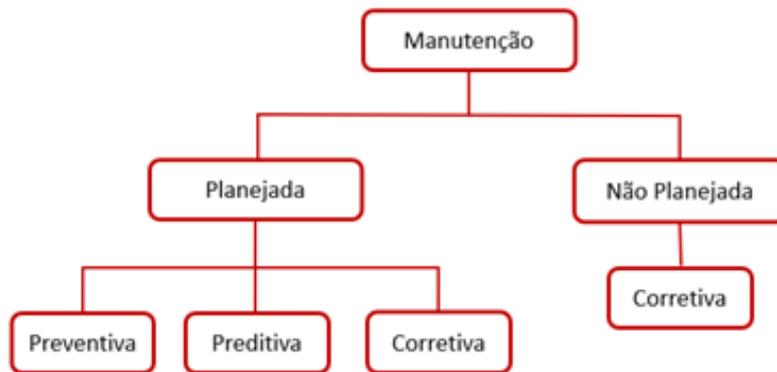
Os dados são levados ao software através *do Internet of Things* (IoT) de forma remota para o software onde é trabalhado o *BI* e em seguida *ML* para preparar para a próxima manutenção.

Mas vem a questão do “O que é Business Intelligence e como ele pode nos ajudar dentro da manutenção preditiva?”, “O que é manutenção preditiva?”, “Há outros tipos de manutenções?”, “Qual a diferença entre as outras manutenções?”, “O que é *Machine Learning*?”, “Como *Machine Learning* pode ajudar?”.

- Tipos de manutenções segundo a NBR-5462, suas vantagens e desvantagens:

Dentro da manutenção temos a manutenção planejada e a não-planejada.

Figura 2 – Diagrama de manutenção



Fonte: Próprio autor

Dentro da manutenção planejada há três tipos de manutenções:

- **Manutenção preditiva:** acompanha os equipamentos/máquinas periodicamente através dos dados que são coletados por meio de inspeções, utilizando as técnicas de análise de vibração, ultrassom, inspeção visual, entre outras técnicas não destrutivas.

**Vantagens:** Reduz a frequência de manutenções corretivas e preventivas, já que seu objetivo é prever as falhas e conseqüentemente como resultado aumentando a vida útil do equipamento e melhorando seu desempenho.

**Desvantagens:** Envolve alto investimento para compra de equipamentos para esse tipo de manutenção (sensores, por exemplo).

- **Manutenção preventiva:** feita periodicamente (intervalos predeterminados) ou de acordo com os critérios prescritos, reduzindo a probabilidade de falha ou degradação do

equipamento. agindo antes da manutenção corretiva que acaba resultando em altos custos, promovendo reparos melhorando as condições dos equipamentos.

**Vantagens:** Reduz a probabilidade de falha ou degradação do equipamento.

**Desvantagens:** Não é a manutenção com o melhor custo/ benefício, tendo seu custo três vezes maior que a manutenção preditiva.

- **Manutenção corretiva:** A manutenção é efetuada após a ocorrência de falha ou quebra, destinada a recolocação do item.

**Vantagens:** Não há.

**Desvantagens:** A manutenção mais cara e tomando um tempo maior, trazendo mais prejuízo para a empresa.

A manutenção não-planejada é somente uma (a manutenção corretiva), onde também a encontramos na manutenção planejada.

Com a evolução tecnológica, quando falamos de tecnologia não estamos falando somente de máquinas, mas também de novos softwares, técnicas e linguagens de programação. Com a implementação do *Business Intelligence* ajudou na tomada de decisão com a ajuda de seus softwares: Tableau, Microsoft Power BI, Google Data Studio, Domo, entre outros através da mineração de dados, visualização de dados, ferramentas/infraestrutura de dados e as práticas recomendadas para auxiliar as organizações na tomada de decisão impulsionadas por dados.

Também podemos destacar a linguagem de programação e suas técnicas como *Machine Learning*, que ajudam da predição de uma determinada necessidade podendo saber o próximo dia que será preciso fazer a manutenção de um determinado equipamento, ou até mesmo prever qual equipamento terá a necessidade de manutenção. Isso é possível através de suas técnicas e vendo qual é a melhor escolha para cada situação.

Esse trabalho tem a proposta de mostrar como o *Business Intelligence* atua junto a manutenção preditiva dentro da indústria 4.0 e suas dificuldades, tornando a manutenção mais eficiente não causando atraso nas produções e com o custo menor em comparação às demais manutenções.

## 1.1 – Business Intelligence e Machine Learning na Manutenção Preditiva

Com o passar dos anos a tecnologia evoluiu rapidamente, por exemplo, a internet 5G, carregar seu celular por indução, reconhecimento facial, equipamentos industriais mais modernos e até novas técnicas de manutenção dando mais segurança à indústria para produção de seus produtos. *Business Intelligence* é conhecido como métodos de coletas e uso de informações para aumentar o fator competitivo. É um conjunto de ferramentas de infraestrutura de dados, que são usados para gerar benefícios e mudanças positivas. *Machine Learning* é conhecido como aprendizado de máquinas onde suas inúmeras técnicas ajudam nas previsões, fazendo com que a máquina pense por ela mesmo até chegar no resultado esperado. Com essas grandes ferramentas tecnológicas mostrando a utilização delas dentro da “Quarta Revolução Industrial” ou também conhecida como Indústria 4.0.

### 1.1.1 – Business Intelligence

A origem do termo *Business Intelligence*, teve sua primeira vez na década de 50 por um pesquisador da IBM, Hans Peter Luhn. Em seu artigo chamado “A *Business Intelligence System*” escrito em 1958 propõe desenvolvimento de sistema automático tendo como base máquinas de processamentos de dados, que indexa e codifica os documentos de forma automática.

Como suas principais características temos:

- Capacidade de extrair e integrar dados de diversas fontes
- Transformar registros obtidos para informação útil direcionado para conhecimento empresarial
- Análise de dados contextualizada
- Trabalhando com hipótese e desenvolvimento de estratégias e ações competitivas

Em 1989 Howard Dresner cunhou uma definição moderno do termo no sentido em que normalmente é usado hoje na indústria. (“acesso do usuário final e análise de conteúdo estruturado, ou seja, dados”)

Segundo a empresa ASAS, *Business Intelligence* é “Uma ampla categoria de programas, aplicações e tecnologias para coletar, armazenar, analisar e fornecer acesso a dados para ajudar os usuários corporativos a tomar as melhores decisões de negócios”...“A

arte de obter uma vantagem comercial a partir de dados, respondendo perguntas fundamentais”.

No final século XX para o início do século XXI, o que antes era um diferencial competitivo, acabou se tornando algo obrigatório, com a diferença de que muitas empresas ainda não conseguem fazer a utilização do conceito de *Business Intelligence*.

De alguns anos para cá o Business Intelligence deu uma evoluída e passou a incluir processos e atividades para ter melhora em seu desempenho. Esses processos são:

**Mineração de dados** – usa-se aprendizado de máquinas revelando tendências e conjuntos de dados grandes, e banco de dados.

**Geração de relatórios** – processo de compartilhamento de análise para tomada de decisões para as partes interessadas.

**Benchmarking e métricas de desempenho** – comparação dos dados atuais e históricos para o acompanhamento do desenvolvimento em relação às metas

**Análise descritiva** – uso da análise dos dados para descobrir o que aconteceu

**Consultas** – entendendo os dados fazendo perguntas específicas

**Análise estatística** – explorando conceitos estatísticos dos dados com mais profundidade

**Visualização de dados** – transformando análises em dados de representações visuais (gráficos, diagramas, histogramas)

**Análise visual** – exploração dos dados através de histórias visuais para comunicar informações conforme necessário e manter o fluxo da análise.

**Preparação de dados** – identificar as dimensões e prepará-las para sua análise.

### 1.1.2 – Machine Learning

Iniciou a partir da década de 1950. Nessa época Alan Turing , conhecido como “o pai da informática moderna” iniciou o teste de o quanto as máquinas seriam capazes de pensar por elas e elaborar raciocínio como seres humanos.

Alguns anos depois, em 1956 o professor de matemática John McCarthy usou o termo Inteligência Artificial em público, sendo usada como um título do projeto no Dartmouth Summer Research on *Artificial Intelligence*.

*Machine Learning* é um conjunto de algoritmos que consiste em uma linha de ação em face de comando. Também considerado como sequencias de passos para realizar uma tarefa, também ajudando a solução de problemas (algoritmos gerando respostas). Por exemplo, alguns algoritmos (como os do Google) tem a função de mostrar resultados considerando uma certa quantidade de informação prévia como dispositivo usado, horário, localização, entre outras informações.

Desde 1956 os algoritmos sofrem alterações e isso sempre vai acontecer, até para melhorar a funcionalidade de acordo com a evolução tecnológica (máquinas, produções, entre outras). Dentro do *Machine Learning* há alguns tipos de aprendizados, supervisionado e o não-supervisionado.

- **Aprendizado supervisionado:** quando falamos de aprendizado supervisionado é quando se tem uma pergunta de negócio, como por exemplo “Qual a ação de marketing mais impacta nas vendas?”. Com base nesse questionamento identificaria as causa que impactam na variável “vendas”. A aprendizagem supervisionada é dividida entre:

- **Regressão:** método que busca encontrar a variável que evolui em relação a outras, esta entre os métodos mais comuns.

- **Classificação:** método que busca explicar a variável categórica, e não numéricas como no caso da regressão.

- **Aprendizado não-supervisionado:** nesse caso é necessário que a própria máquina faça a exploração das informações localizando os resultados adequados para a pesquisa, não existindo uma resposta especificada ser buscada e sim uma análise que tem resultado nas alternativas para sustentar uma determinada ação. Podem alterar os resultados conforme as variáveis consideradas.

- **Clusterização:** uma técnica que permite dividir automaticamente o conjunto de dados em grupo de acordo com uma similaridade.

- **Deteção de anomalias:** faz a descoberta automaticamente de pontos incomuns no conjunto de dados, util para a identificação de transações fraldulentas, na descoberta de peças defeituosas de hardware e também na identificação de erro humano na entrada de informação.

- **Mineração de associação:** faz a identificação de um conjunto de itens que com frequencia ocorre junto com seu conjunto de dados. Essa análise costuma a ser usada em carrinhos de compras, analisam quais itens costumam ser comprados juntos, desenvolvendo a estratégia mais eficaz de marketing e merchandising.

- **Modelos de variantes latentes:** usados no pré processamento de dados, reduzindo número de recursos em um conjunto de dados.

Em resumo, o aprendizado de máquina não supervisionado descobre padrões desconhecidos por dados. Se é desconhecido não se sabe qual deve ser o resultado, tornando mais aplicável aos problemas do mundo real.

## 1.2 – Justificativa

Com a utilização do *Business Intelligence* dentro da indústria 4.0 houve melhorias em suas produções, não precisando interrompê-las por um tempo maior para fazer reparos, sendo feita com previsão facilitando saber qual peça necessária de troca sempre mantendo e melhorando a vida útil do equipamento e também levando em consideração o aumento do lucro da empresa. A importância dessa pesquisa exploratória que é estudar os dados que a base de dados te dá para que consiga ter bons resultados em cima desses dados.

## 1.3 – Objetivos

Mostrar como funciona o *Business Intelligence* e o *Machine Learning* dentro da manutenção preditiva, técnicas a serem usadas, softwares, linguagens de programação que podem ser utilizadas e como escolhê-las e também mostrando como é uma manutenção preditiva sem o uso de *Business Intelligence* e com o uso de *BI*. Implementando a inteligência artificial prevendo falhas, evitando paradas de linha de fabricação reduzindo custos com a manutenção.

## 1.4 – Delimitação

O foco deste trabalho são as indústrias, sejam elas de pequeno, médio ou grande porte. Hoje na indústria 4.0 a parte de maior importância é a manutenção, com o investimento correto a indústria que utiliza o BI em suas manutenções tem um aumento de produção maior e conseqüentemente o aumento do lucro.

Business Intelligence, como falado anteriormente são um conjunto de técnicas, que com sua formação é possível fazer levantamentos financeiros, levantamentos de produtos

mais vendidos, qual o público melhor para venda, localidade, previsões para alguns anos fazendo assim a melhor tomada de decisão para seu crescimento.

## **1.5 – Metodologia**

Através da nossa pesquisa exploratória, colocando em prática nosso conhecimento adquirido dentro de sala de aula, acrescentando com as pesquisas feitas através de artigos que foram compartilhados através do orientador, livros, entre outros trabalhos colocamos em prática as técnicas de *BI*. O experimento foi feito inteiramente pelo autor com suporte do orientador. A estratégia utilizada no trabalho foi procurar uma base de dados retirada de um repositório de *Machine Learning*, onde os dados são perto da realidade da indústria para que possamos utilizar as técnicas de *BI* baseadas em seus dados e explicando o porquê de cada biblioteca, gráficos e mapas utilizados.

## **1.6 – Descrição**

Este trabalho está estruturado da seguinte forma:

No capítulo 2 será falado sobre o embasamento teórico, onde mostraremos através de gráficos a importância da manutenção preditiva, citações de autores

O capítulo 3 será apresentado a parte prática do nosso experimento desde o uso do software até a inteligência artificial junto com as propostas tecnológicas falando um pouco das técnicas que foram utilizadas

Por último chegaremos à conclusão e em trabalhos futuros para essa base de dados que utilizamos no nosso estudo de caso.

# CAPÍTULO 2

## Embasamento Teórico

Nesse capítulo falaremos um pouco sobre a manutenção, a curva PF, a qual é considerada a curva muito importante dentro das manutenções e quando sabemos como utilizá-la será bastante útil podendo ser utilizada também na relação do custo e manutenção. A partir desse momento entendemos qual manutenção vale mais a pena a ser implementada na indústria.

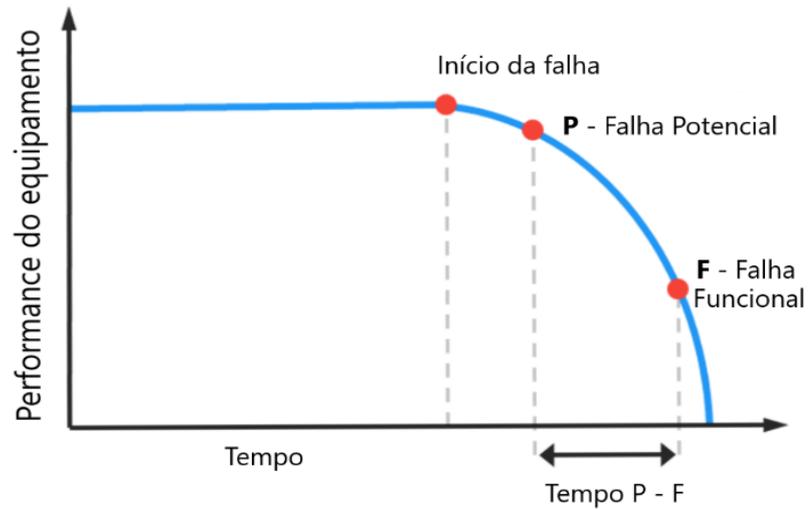
“A primeira revolução industrial durou 201 anos (1712-1913), cedendo lugar à segunda revolução industrial, quando Henry Ford criou a linha de produção em massa, fazendo a produção empurrada, criando o conceito de produção em escala, reduzindo custos e popularizando o produto para que a massa trabalhadora pudesse adquiri-lo[...]então, a era da automação, também conhecida como Terceira Revolução Industrial, com a implantação de computadores no chão da fábrica, controles eletrônicos, sensores e dispositivos capazes de gerenciar uma grande quantidade de variáveis de produção, permitindo a tomada de decisões de controle de dispositivos de maneira autônoma.”(ALMEIDA; PAULO, 2019)

Desde o aparecimento da primeira máquina industrial a vapor na Inglaterra a partir da segunda metade do séc XVIII a manutenção era feita sempre quando uma peça quebrava. Então era necessário parar a produção para fazer a troca da mesma, feito isso a produção retornaria a seus afazeres.

- **Curva PF**

Quando falamos de manutenção, não devemos esquecer de falar sobre a curva PF. A curva PF é uma representação gráfica da vida útil do um equipamento, portanto se conhecer a funcionalidade da curva PF também conhece a vida útil do equipamento e seus pontos de falhas. A curva é chamada da seguinte maneira pelo ponto P significar **performance** e o ponto F significar **falha funcional**.

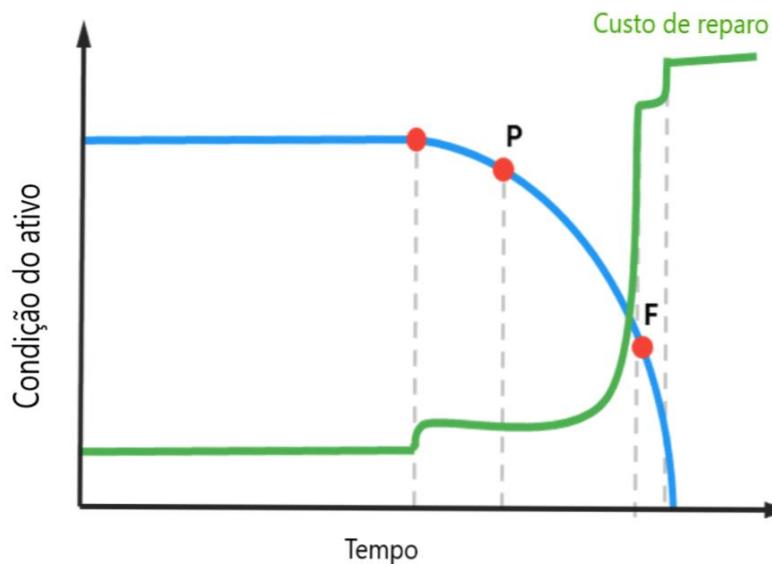
Figura 3 – Curva PF



Fonte: <https://www.power-mi.com/node/43038>

Conforme o tempo vai passando e o equipamento está sendo utilizado, a performance vai caindo e antes disso acontecer é quando a manutenção preditiva deve entrar em cena. O BI identifica a falha antes dela acontecer para que a manutenção seja feita sem afetar a produção e para que a performance do equipamento continue com excelência. Conforme mostra no gráfico abaixo, quanto mais se demora para fazer a manutenção, maior é o custo do reparo da máquina.

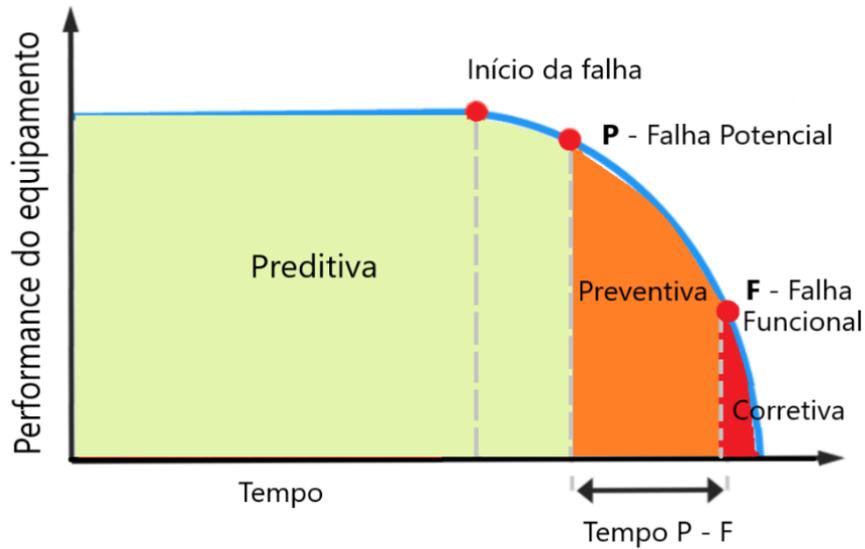
Figura 4 – Curva TxC



Fonte: <https://www.power-mi.com/node/43038>

No próximo gráfico mostra que até o ponto P é onde consideramos que deve ser feita a manutenção preditiva. Em seguida, passando o ponto P até o ponto F é considerado a manutenção preventiva e a partir do ponto F para baixo, é a manutenção corretiva. Chegando no ponto final (falha funcional), é onde mostra a degradação da peça/máquina.

Figura 5 – Manutenções



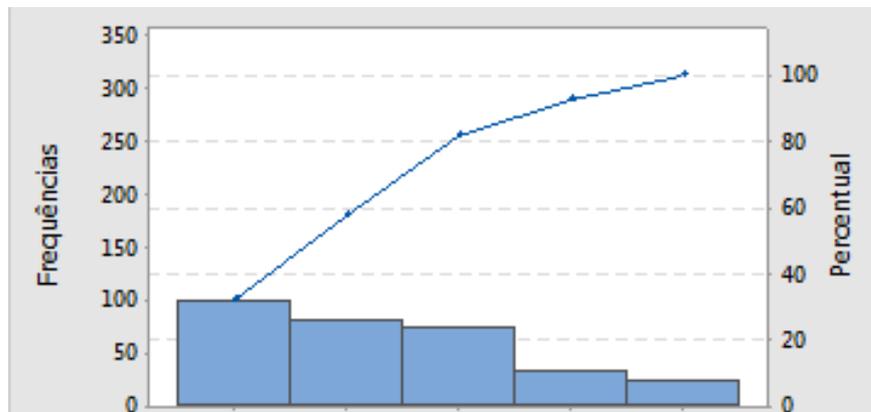
Fonte: <https://www.power-mi.com/node/43038>

- **Pareto**

Pareto formou-se em matemática, física e engenharia na Itália e foi responsável pela formulação da lei de distribuição de renda. Tentou mostrar que a renda e riqueza não eram distribuídas de maneira uniforme ao longo da evolução das sociedades, chamada Lei de Pareto.

O diagrama de Pareto é uma ferramenta de qualidade importante dentro das empresas identificando seus problemas. O diagrama de Pareto é composto por gráficos de barras que ordena as frequências das ocorrências ordenada de forma crescente, permitindo a localização de problemas vitais e eliminação de futuras perdas. Também é conhecida como a ferramenta 80/20, onde 80% são seus resultados e 20% são suas causas.

Figura 6 – Diagrama de Pareto



Fonte : <https://support.minitab.com/>

A terceira revolução industrial durou até 2010. Como se pode ver os intervalos vem diminuindo, inaugurando uma nova era [...] maior protagonista é a internet que já esta consolidada como um grande canal de comunicação convergente de todas as tecnologias, agora sendo colocado dentro da indústria com seus conceitos adaptados a máquinas e equipamentos. (ALMEIDA; PAULO, 2019)

Volpato, Rufino e Dias (2014, p. 5) citam que:

Tão importante quanto gerar informação é a capacidade de processamento de dados volumosos em alta velocidade. Isso se comprova pelo fato de que, nas últimas décadas, presenciamos o desenvolvimento de supercomputadores que atendam essa necessidade: quanto mais a tecnologia foi penetrando no meio social, mais informações as pessoas foram gerando e consumindo.

## 2.1 – Fundamentos da Indústria 4.0

Com a evolução tecnológica e a integração dos processos segundo o conceito da Indústria 4.0, os sistemas de produção passaram a ficar cada vez mais inteligentes, capazes de detectar o surgimento de necessidades produtivas, de suprimentos e de matéria-prima, o que envolve a união de tecnologias físicas e digitais e a integração de todas as etapas do desenvolvimento de um produto ou processo. Isso trouxe um impacto muito positivo, traduzido em maior eficiência e produtividade. (ALMEIDA; PAULO, 2019)

# CAPÍTULO 3

## Experimento e Propostas Tecnológicas

Durante esse capítulo será apresentado as informações sobre o dataset escolhido para o experimento, também abordando um pouco sobre cada biblioteca que foi usada durante toda a nossa prática até o aprendizado de máquinas.

A proposta deste trabalho é apresentar a implementação do BI dentro da manutenção industrial. Mostrando como isso é possível e quais técnicas foram utilizadas no processo. Escolhemos usar um dataset que tivesse dados próximos a realidade que as indústrias passam para manutenção preditiva, o dataset foi retirado do site para trabalhos acadêmicos UCI Machine Learning Repository. O link do dataset se encontra em “Referências Bibliográficas” caso tenham interesse em estudá-lo ou até mesmo ver outros datasets.

Há informações que não sabemos como: tipo de máquina que estamos lidando, produção da máquina e data. As demais informações estabelecidas do dataset são:

- O conjunto de dados consiste em 10.000 linhas de dados armazenados com 14 colunas.
- UID: identificador único variando de 1 a 10.000
- ID do produto: letras L (low 50% de todos os produtos), M (medium 30%) ou H (high 20%) como variantes de qualidade do produto e um número de série específico da variante
- Temperatura do ar [K]: gerado usando um processo de passeio aleatório posteriormente normalizado para um desvio padrão de 2 K em torno de 300 K
- Temperatura do processo [K]: gerado usando um processo de passeio aleatório normalizado para um desvio padrão de 1 K, adicionado à temperatura do ar mais 10 K.
- Velocidade de rotação [rpm]: calculada a partir de uma potência de 2860 W, sobreposta com um ruído normalmente distribuído
- Torque [Nm]: os valores de torque são normalmente distribuídos em torno de 40 Nm com um  $\ddot{f} = 10$  Nm e nenhum valor negativo.
- Desgaste da ferramenta [min]: As variantes de qualidade H / M / L adicionam 5/3/2 minutos de desgaste da ferramenta à ferramenta usada no processo.

O rótulo de 'falha da máquina' que indica se a máquina falhou neste ponto de dados específico para qualquer um dos seguintes modos de falha é verdadeiro.

A falha da máquina consiste em cinco modos de falha independentes

- Falha por desgaste da ferramenta (TWF): a ferramenta será substituída ou falha em um tempo de desgaste da ferramenta selecionado aleatoriamente entre 200 - 240 minutos (120 vezes em nosso conjunto de dados).

- Falha de dissipação de calor (HDF): a dissipação de calor causa uma falha de processo, se a diferença entre a temperatura do ar e do processo for inferior a 8,6 K e a velocidade de rotação da ferramenta for inferior a 1380 rpm. Esse é o caso de 115 pontos de dados.

- Falha de energia (PWF): o produto do torque e da velocidade de rotação (em rad / s) é igual à potência necessária para o processo. Se esta potência estiver abaixo de 3500 W ou acima de 9000 W, o processo falha, o que é o caso 95 vezes em nosso conjunto de dados.

- Falha de esforço excessivo (OSF): se o produto do desgaste da ferramenta e torque exceder 11.000 minNm para a variante de produto L (12.000 M, 13.000 H), o processo falha devido ao esforço excessivo. Isso é verdade para 98 pontos de dados.

- Falhas aleatórias (RNF): cada processo tem uma chance de 0,1% de falhar independentemente de seus parâmetros de processo. Este é o caso de apenas 5 pontos de dados, menos do que poderia ser esperado para 10.000 pontos de dados em nosso conjunto de dados.

Para esse trabalho foi utilizado o software Anaconda, podendo utilizar a linguagem Python ou linguagem R, sendo a plataforma de código aberto mais popular para ciência de dados Python. Dentro do software há pacotes para serem instalados e gerenciados com facilidade conforme sua necessidade, nesse caso instalamos o Jupyter Notebook onde ele já vem com Python.

Importante ressaltar que há outros softwares que podem ser utilizados, como por exemplo o Google Colab, como também há outras linguagens de programação que possam ser utilizadas como linguagem R, Julia, Java, entre outras. O link do colab encontra-se dentro de "Referências Bibliográficas" caso tenham interesse de ver o desenvolvimento do trabalho.

A seguir será mostrado o passo a passo do trabalho:

### Passo 1 :

Iniciamos instalando três bibliotecas: *matplotlib*, *seaborn* e *scikit-learn*.

Figura 7 – Instalação das bibliotecas iniciais

```
In [1]: !pip install matplotlib
!pip install seaborn
!pip install scikit-learn
```

Fonte: Próprio autor

Foram utilizadas as bibliotecas acima por causa da sua funcionalidade:

**Matplotlib** – biblioteca para criação de gráficos

**Seaborn** – biblioteca de visualização de dados baseada no matplotlib, porém com gráficos mais sofisticados.

**Scikit-learn** – biblioteca para Inteligência Artificial

Instalado, receberá mensagens de instalação e depois retornando que a instalação esta completa. Em seguida precisamos fazer com que o Jupyter Notebook leia o nosso arquivo em formato é .csv.

### Passo 2 :

Figura 8 – Lendo o arquivo csv

```
In [2]: import pandas as pd
dados = pd.read_csv("ai4i2020.csv")
display(dados)
```

Fonte: Próprio autor

Para que nosso arquivo possa ser lido, precisaremos importar a biblioteca pandas. O Pandas é uma biblioteca de manipulação, leitura e visualização de dados. Nesse caso usaremos para leitura. Utilizamos o “display” por usarmos um DataFrame (df) onde foi substituído para o nome de “dados”.

O resultado disso é:

Tabela 1 – Resultado do DataFrame

	UDI	Product ID	Type	Air temperature [K]	Process temperature [K]	Rotational speed [rpm]	Torque [Nm]	Tool wear [min]	Machine failure	TWF	HDF	PWF	OSF	RNF
0	1	M14860	M	298.1	308.6	1551	42.8	0	0	0	0	0	0	0
1	2	L47181	L	298.2	308.7	1408	46.3	3	0	0	0	0	0	0
2	3	L47182	L	298.1	308.5	1498	49.4	5	0	0	0	0	0	0
3	4	L47183	L	298.2	308.6	1433	39.5	7	0	0	0	0	0	0
4	5	L47184	L	298.2	308.7	1408	40.0	9	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
9995	9996	M24855	M	298.8	308.4	1604	29.5	14	0	0	0	0	0	0
9996	9997	H39410	H	298.9	308.4	1632	31.8	17	0	0	0	0	0	0
9997	9998	M24857	M	299.0	308.6	1645	33.4	22	0	0	0	0	0	0
9998	9999	H39412	H	299.0	308.7	1408	48.5	25	0	0	0	0	0	0
9999	10000	M24859	M	299.0	308.7	1500	40.2	30	0	0	0	0	0	0

Fonte: Próprio autor

Observamos que nosso DataFrame tem 10.000 linhas e 14 colunas, para ter certeza se esta tudo correto usamos o `print(dados.info())` onde o programa entende que queremos imprimir toda a informação do meu arquivo “dados”.

### Passo 3 :

Figura 9 – Imprimindo informações do DataFrame

```
In [3]: print(dados.info())

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   UDI                                    10000 non-null  int64
 1   Product ID                            10000 non-null  object
 2   Type                                   10000 non-null  object
 3   Air temperature [K]                   10000 non-null  float64
 4   Process temperature [K]                10000 non-null  float64
 5   Rotational speed [rpm]                 10000 non-null  int64
 6   Torque [Nm]                            10000 non-null  float64
 7   Tool wear [min]                        10000 non-null  int64
 8   Machine failure                        10000 non-null  int64
 9   TWF                                    10000 non-null  int64
10  HDF                                    10000 non-null  int64
11  PWF                                    10000 non-null  int64
12  OSF                                    10000 non-null  int64
13  RNF                                    10000 non-null  int64
dtypes: float64(3), int64(9), object(2)
memory usage: 1.1+ MB
None
```

Fonte: Próprio autor

#### Passo 4 :

Esse início é importante para sabermos como iremos tratar os dados, se tivesse alguma linha faltando informação, precisaríamos tratá-las. Por opção do autor, foi decidido renomear as colunas (`dados.rename(columns={...})`).

Figura 10 – Renomeando as colunas do DataFrame

```
In [5]: dados.rename(columns={'Product ID':'ID do Produto', 'Type':'Tipo', 'Air temperature [K]':'Temperatura ambiente [K]',
                             'Process temperature [K]':'Temperatura do processo [K]',
                             'Rotational speed [rpm]':'Velocidade rotacional [rpm]',
                             'Tool wear [min]': 'Desgaste da ferramenta [min]', 'Machine failure' : 'Falha técnica'}, inplace=True)
dados
```

Out[5]:

	UDI	ID do Produto	Tipo	Temperatura ambiente [K]	Temperatura do processo [K]	Velocidade rotacional [rpm]	Torque [Nm]	Desgaste da ferramenta [min]	Falha técnica	TWF	HDF	PWF	OSF	RNF
0	1	M14860	M	298.1	308.6	1551	42.8	0	0	0	0	0	0	0
1	2	L47181	L	298.2	308.7	1408	46.3	3	0	0	0	0	0	0

Fonte: Próprio autor

#### Passo 5 :

Iniciando trabalho com gráficos:

Figura 11 – Importando outras bibliotecas

```
In [6]: import seaborn as sns
import matplotlib.pyplot as plt

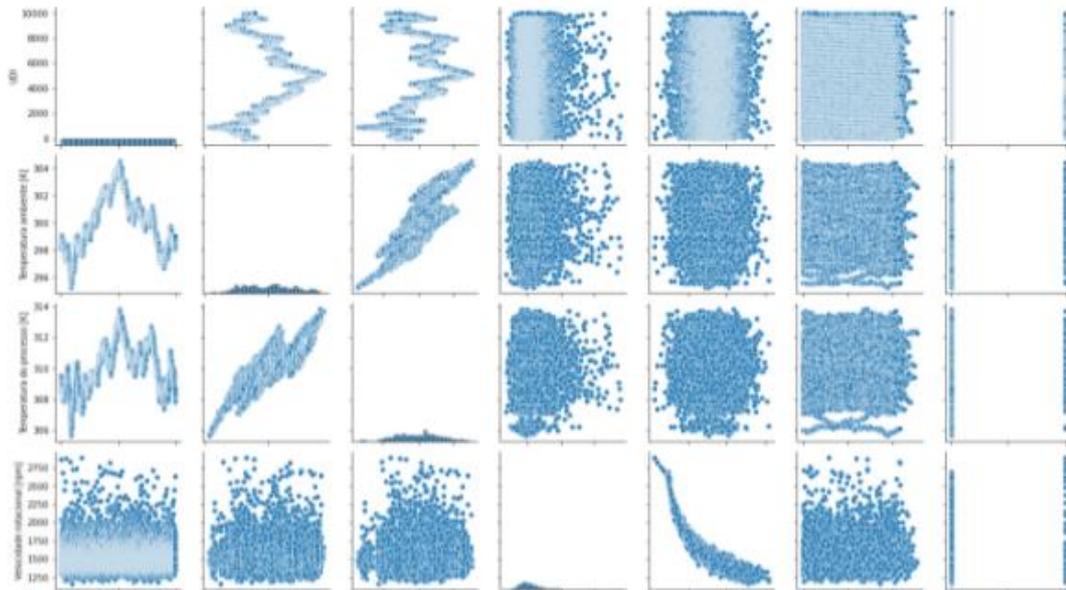
sns.pairplot(dados)
plt.show()
```

Fonte: Próprio autor

Nesse momento importamos o *sns* e o *plt*. Ou seja, criamos gráficos com o *seaborn* e mostramos o gráfico com *matplotlib*. O *pairplot* do *sns* cria diferentes gráficos onde ocorre a correlação dentro da tabela.

Resultando em :

Figura 12 – Correlação existentes na tabela



Fonte: Próprio autor

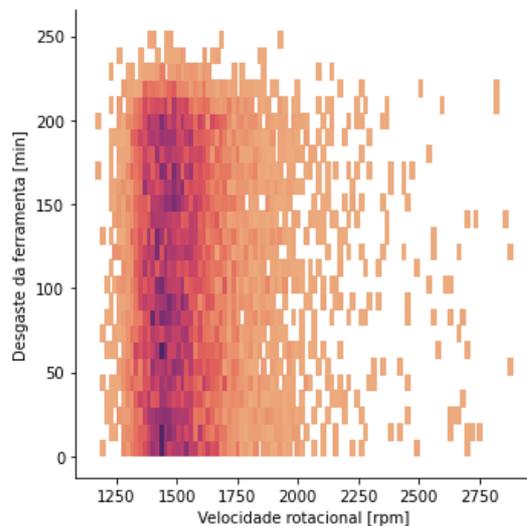
O próximo gráfico que fizemos da biblioteca *sns* é o *displot*. Um gráfico de distribuição, onde a cor mais forte é onde eles se relacionam mais.

### Passo 6 :

Figura 13 – Gráfico de distribuição

```
In [25]: sns.displot(dados, x="Velocidade rotacional [rpm]",y="Desgaste da ferramenta [min]",  
                  cmap = 'flare' )
```

```
Out[25]: <seaborn.axisgrid.FacetGrid at 0x1d574e9d340>
```

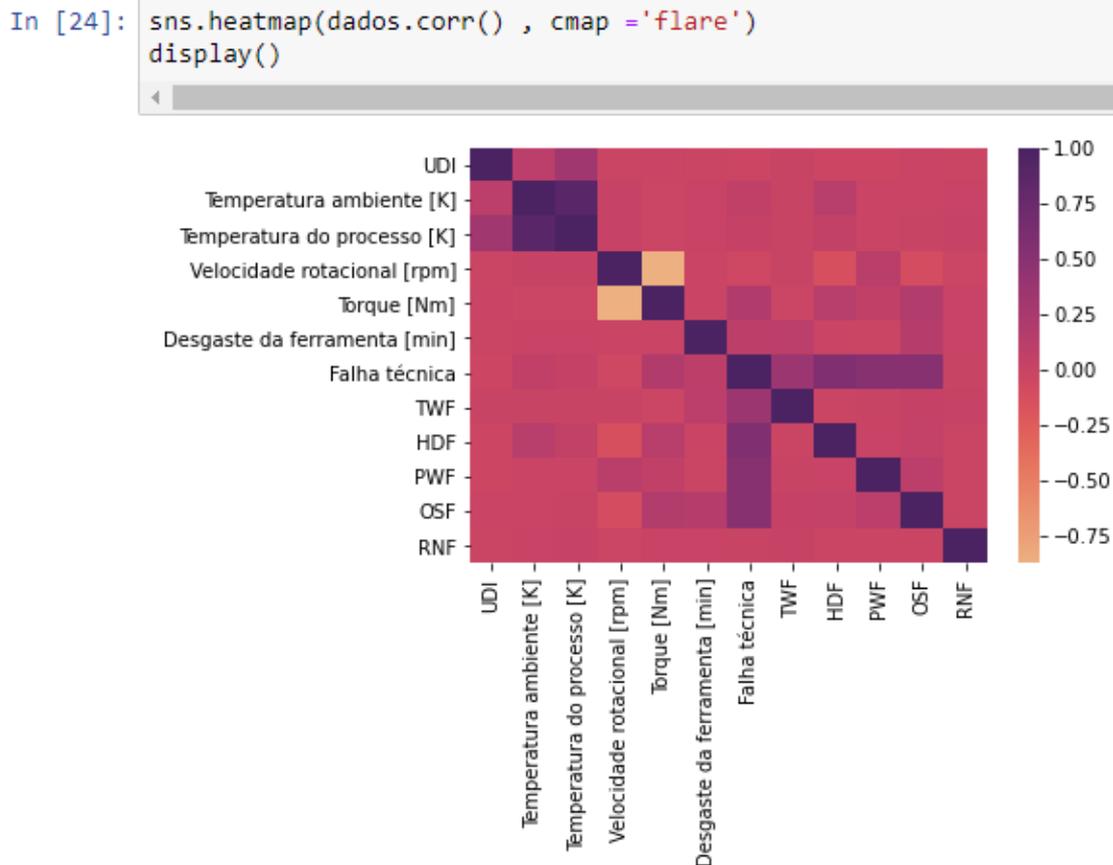


Fonte: Próprio autor

A correlação é vista da melhor maneira através do mapa de calor, conhecido também como *heatmap*. A correlação é vista quando o mapa de calor mostra a correlação dentro da tabela, sua correlação é forte quando for perto de 1,0 ou 1,0.

### Passo 7 :

Figura 14 – Gráfico de calor



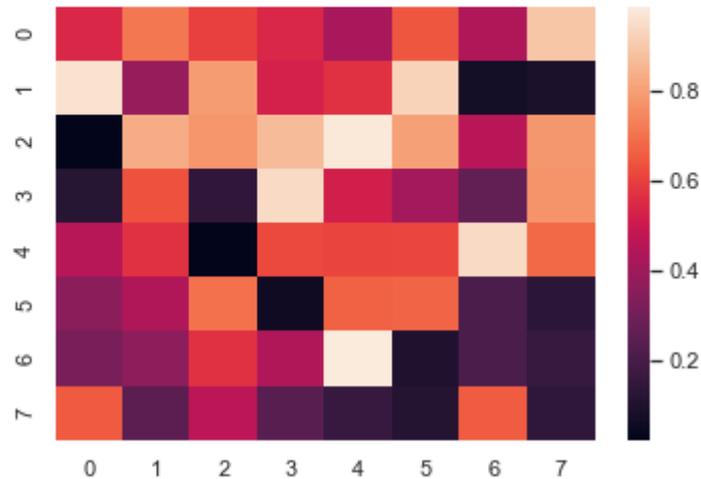
Fonte: Próprio autor

Fizemos um gráfico de sementes aleatórias, conhecido como *random.seed* (0) declarando *nprand* para que não tenha modificação em nosso dataset. Utilizamos 0 (zero) tornando os números aleatórios previsíveis.

## Passo 8 :

Figura 15 – Gráfico random seed

```
In [9]: import numpy as np; np.random.seed(0)
import seaborn as sns; sns.set_theme()
nprand = np.random.rand(8,8)
ax = sns.heatmap(nprand)
```



Fonte: Próprio autor

Agora partimos para a inteligência artificial, onde utilizaremos o *numpy* para informar as variáveis. O *y* é o dado que eu quero prever e escolhemos prever através da Falha Técnica, e o *X* escolhemos utilizar a HDF (falha de dissipação de calor) e OSF (falha de esforço excessivo). Em seguida separaremos os dados de treino e teste e a biblioteca a ser utilizada é o *sklearn*. A ordem sempre será *x\_treino*, *x\_teste*, *y\_treino*, *y\_teste*, caso for trocada uma ordem dará erro no processo.

Quando houver treino e teste existe um padrão, dentro do *train\_test\_split* usa 30% dos dados para fazer o teste e o restante dos dados (70%) é utilizado para treino, isso pode ser modificado, mas nesse caso não há necessidade.

Escolhemos testar a treinar a regressão linear e árvore de decisão, importamos *LinearRegression* e *RandomForestRegressor*. O próximo passo é criar a inteligência, então chamamos o *modelo\_regressao linear* e *modelo\_arvoredecisao* para criarmos o treinamento com o *modelo\_regressao linear.fit(x\_treino, y\_treino)*. O resultado retorna com a técnica que melhor se encaixa, sendo nesse caso sendo *RandomForestRegressor*.

## Passo 9 :

Figura 16 – Organização dos dados para treino e teste

```
In [25]: import numpy as np

y = dados ["Falha técnica"]
x = dados [["HDF", "OSF"]]

#separar os dados de treino e teste
from sklearn.model_selection import train_test_split
x_treino, x_teste, y_treino, y_teste = train_test_split
(x, y, test_size=0.3)

#regressao linear e arvore de decisao
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor

#cria a inteligencia
modelo_regressaolinear = LinearRegression()
modelo_arvoredecisao = RandomForestRegressor()

#treina a inteligencia
modelo_regressaolinear.fit(x_treino, y_treino)
modelo_arvoredecisao.fit(x_treino, y_treino)
```

Out[25]: RandomForestRegressor()

Fonte: Próprio autor

Em seguida importamos `r2_score` da biblioteca `sklearn` para determinar a pontuação de regressão, a melhor pontuação possível é 1,0 ou próximo.

## Passo 10 :

Figura 17 - Resultados

```
In [26]: from sklearn.metrics import r2_score

#cria as previsões
previsao_regressaolinear = modelo_regressaolinear.predict(x_teste)
previsao_arvoredecisao=modelo_arvoredecisao.predict(x_teste)

#compara com o gabarito

print(r2_score(y_teste, previsao_regressaolinear))
print(r2_score(y_teste, previsao_arvoredecisao))

0.6106780639284046
0.6210796356071233
```

Fonte: Próprio autor

Conforme mencionado no primeiro capítulo, um dos desafios das indústrias é implementar o uso do BI dentro delas, isso envolve custos, equipe especializada e equipamentos corretos. E o desafio que se encontra dentro do BI é saber tratar os dados de forma correta sem saber qual técnica pode ser usada para cada situação. Pode-se dizer que alguns momentos são feitos tentativa e erro com as técnicas mais prováveis de dar certo.

# CAPÍTULO 4

## Conclusão e Trabalhos Futuros

O presente estudo de caso em que foi trabalhado propôs a análise de possíveis falhas existentes na manutenção industrial dos métodos de BI dentro da indústria 4.0 com o objetivo de solucionar problema antes das tais ocorrências, não colocando em risco os procedimentos produtivos e financeiros.

Conforme apontados durante o estudo, os métodos empregados estão voltados a aplicação de conceitos referentes a sistemas de gestão e qualidade em manutenção preditiva, engenharia e sustentabilidade.

O método pesquisado deixou caminho para trabalhos futuros para aperfeiçoar o resultado final dataset. Nesse sentido é válido destacar os experimentos pertinentes aos estudos realizados pela análise computacional neste estudo de caso, aplicado na manutenção preditiva em BI.

A análise do nosso estudo de caso permitiu concluir que nosso objetivo foi concluído ao perceber que foi possível aplicar as técnicas de BI nesse estudo, antes de uma parada parcial do equipamento. Um novo modelo de indústria surge com o aprimoramento da manutenção preditiva através das aplicações computacionais, solucionando problemas preditivos.

# Referências Bibliográficas

**Machine Learning Repository**. Disponível em: <<https://archive.ics.uci.edu/ml/datasets/AI4I+2020+Predictive+Maintenance+Datas+et.>>. Acesso em: 18/08/2021.

SOUSA, Rayane. **EXPERIMENTO DATASET AI4I2020**. Disponível em: <<https://colab.research.google.com/drive/1FDPBqnnhSRw1Evqp-WexkBJX0LBIycg3?usp=sharing>>. Criado em: 19/08/2021

VOLPATO Tiago. RUFINO, Ricardo Ribeiro. DIAS, Jaime William. **BIG DATA – TRANSFORMANDO DADOS EM DECISÕES**. UNIPAR. Disponível em: <<https://docplayer.com.br/3437738-Big-data-transformando-dados-em-decisoes.html>> Acesso em: 11/09/2021.

SANTOS, P, Alexandre. POZSETTI, T, João Vitor, MORAES, V, A, Priscilla. AVELINO, Cleide Henrique. **Utilização da ferramenta Diagrama de Pareto para auxiliar na identificação dos principais problemas nas empresas**. Disponível em: <<https://unisalesiano.com.br/aracatuba/wp-content/uploads/2020/12/Artigo-Utilizacao-da-ferramenta-Diagrama-de-Pareto-para-auxiliar-na-identificacao-dos-principais-problemas-nas-empresas-Pronto.pdf>>. Acesso em: 15/10/2021.

STELLA, Jefferson. **Tudo sobre Big Data: a revolução da análise de dados**. Acesso em: 20/10/2021

Ławrynowicz, A.; Tresp, V. **Introducing machine learning. In Perspectives on Ontology Learning**; Lehmann, J., Voelker, J., Eds.; IOS Press: Heidelberg, Germany, 2014. Acesso em: 01/11/2021

**INDUSTRIA 4.0**. Disponível em: <https://www.portaldaindustria.com.br/industria-de-a-z/industria-4-0/>. Acesso em: 08/11/2021

RODRIGUES, Juliano. **Business Intelligence: Inteligência nos Negócios**. Disponível em: <[https://d1wqtxts1xzle7.cloudfront.net/33586058/Fabiana\\_Tronto-with-cover-page-v2.pdf?Expires=1640224448&Signature=ShoS8STx34ehMMLK3AsJBgw41xx6HLQnKMeb0YkRdTW9I99YIeOFjaY1KmwWJbtFA0--va6A87MwyEB9IA9nOGMIv8VyfWJ0d6FWSCVIXxpXwmvI5Lv6s5cnm7tVITLs2zDl~jtd2dMo9t4qoP7M7FBNbTNjSmIcL0OoRWQx8tGPMduD7WZM2PeIMPEVTOX-lcxg7mgravEnap8s~MEta-uNdL5SbnJe6QSJhhtE96g50ZtYR99bo8590p4xfziOuj1Ov-KL4Q0kJKwM~CaNntjD66176z83anNc5KElWsDI7avieFjTweysOLLotTZ-kXIIx-YI09X6m7e9FxkcNw\\_\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzle7.cloudfront.net/33586058/Fabiana_Tronto-with-cover-page-v2.pdf?Expires=1640224448&Signature=ShoS8STx34ehMMLK3AsJBgw41xx6HLQnKMeb0YkRdTW9I99YIeOFjaY1KmwWJbtFA0--va6A87MwyEB9IA9nOGMIv8VyfWJ0d6FWSCVIXxpXwmvI5Lv6s5cnm7tVITLs2zDl~jtd2dMo9t4qoP7M7FBNbTNjSmIcL0OoRWQx8tGPMduD7WZM2PeIMPEVTOX-lcxg7mgravEnap8s~MEta-uNdL5SbnJe6QSJhhtE96g50ZtYR99bo8590p4xfziOuj1Ov-KL4Q0kJKwM~CaNntjD66176z83anNc5KElWsDI7avieFjTweysOLLotTZ-kXIIx-YI09X6m7e9FxkcNw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA)>. Acesso em: 15/11/2021.

L. C. Nepomuceno. **Técnicas de Manutenção Preditiva**. Editora Edgard Blucher, São Paulo, 2012.

**LA CURVA P-F**. Disponível em: <https://www.power-mi.com/node/43038>. Acesso em : 22/11/2021

**BEM VINDOS A QUARTA REVOLUÇÃO INDUSTRIAL**. Disponível em: <http://especiais.estadao.com.br/mundodigital/ola-mundo/>. Acesso em: 05/12/2021.

**PROFESSOR JOHN MACCARTHY**. Disponível em: <<http://jmc.stanford.edu/>>. Acesso em: 20/12/2021

**O QUE É BUSINESS INTELLIGENCE?**. Disponível em: <<https://www.tableau.com/pt-br/learn/articles/business-intelligence>>. Acesso em: 25/11/2021.

LUHN, Hans Peter. **A Business Intelligence System**. Disponível em: <<http://altaplana.com/ibm-luhn58-BusinessIntelligence.pdf>>. Acesso em: 24/12/2021

ELLIOT, Timo. **Happy Birthday to the “Father of Business Intelligence”**. Disponível em: <<https://blogs.sap.com/2013/07/01/happy-birthday-to-the-father-of-business-intelligence/>>. Acesso em: 23/11/2021.

# Apêndice I

Artigo – Hans Peter Luhn “A Business Intelligence System”

H. P. Luhn

## A Business Intelligence System

**Abstract:** An automatic system is being developed to disseminate information to the various sections of any industrial, scientific or government organization. This intelligence system will utilize data-processing machines for auto-abstracting and auto-encoding of documents and for creating interest profiles for each of the “action points” in an organization. Both incoming and internally generated documents are automatically abstracted, characterized by a word pattern, and sent automatically to appropriate action points. This paper shows the flexibility of such a system in identifying known information, in finding who needs to know it and in disseminating it efficiently either in abstract form or as a complete document.

### Introduction

Efficient communication is a key to progress in all fields of human endeavor. It has become evident in recent years that present communication methods are totally inadequate for future requirements. Information is now being generated and utilized at an ever-increasing rate because of the accelerated pace and scope of human activities and the steady rise in the average level of education. At the same time the growth of organizations and increased specialization and divisionalization have created new barriers to the flow of information. There is also a growing need for more prompt decisions at levels of responsibility far below those customary in the past. Undoubtedly the most formidable communications problem is the sheer bulk of information that has to be dealt with. In view of the present growth trends, automation appears to offer the most efficient methods for retrieval and dissemination of this information.

During the past decade significant progress has been made in applying machines to the processes of information retrieval. Automatic dissemination has so far been given little consideration; however, unless substantial portions of human effort in this area can be replaced by automatic operations, no significant over-all improvement will be achieved. Even the information retrieval processes mechanized so far still require appreciable human effort to organize the information before it is entered into machines.

It is believed that techniques now being developed will greatly contribute to the solution of the problem by extending automatic processes to the preparatory phases of mechanical information-retrieval systems, to the area of dissemination and to associated functions. Ideally, an automatic system is needed which can accept information

in its original form, disseminate the data promptly to the proper places and furnish information on demand.

The techniques proposed here to make these things possible are:

1. Auto-abstracting of documents;
2. Auto-encoding of documents;
3. Automatic creation and updating of *action-point* profiles.

All of these techniques are based on statistical procedures which can be performed on present-day data processing machines. Together with proper communication facilities and input-output equipment a comprehensive system may be assembled to accommodate all information problems of an organization. We call this a *Business Intelligence System*.

### Objectives and principles

Before the system operation is described, the term *Business Intelligence System* should be defined and the objectives and principles stated.

In this paper, *business* is a collection of activities carried on for whatever purpose, be it science, technology, commerce, industry, law, government, defense, et cetera. The communication facility serving the conduct of a business (in the broad sense) may be referred to as an *intelligence system*. The notion of *intelligence* is also defined here, in a more general sense, as “the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal.”<sup>1</sup>

The term *document* is used to designate a block of information confined physically in a medium such as a

letter, report, paper or book. The term may also include the medium itself.

The objective of the system is to supply suitable information to support specific activities carried out by individuals, groups, departments, divisions, or even larger units. These are the *action points* previously referred to. To this end the system concerns itself with the admission or acquisition of new information, its dissemination, storage, retrieval and transmittal to the action points it serves.

More particularly the object of the system is to perform these functions speedily and efficiently, taking advantage of novel procedures which utilize the inherent capabilities of electronic devices.

One of the most crucial problems in communication is that of channeling a given item of information to those who need to know it. Present methods of accomplishing this are inadequate and the general practice is to disseminate information rather broadly to be on the safe side. Since this method tends to swamp the recipients with paper, the probability of not communicating at all becomes great. The Business Intelligence System provides means for selective dissemination to each of its action points in accordance with their current requirements or desires. This is accomplished by the mechanical creation of *profiles* reflecting the sphere of interest of each point and by updating these profiles as dictated by changes in the attitude of the respective action points and as recorded by the system on the basis of certain transactions.

Another problem in communication is to discover the person or section within an organization whose interests or activities coincide most closely with a given situation. Presently, the difficulty of finding such relationships often results in improper decisions, wrong actions, inaction, or duplication. An objective of the Business Intelligence System is to identify related interests by use of profiles of action points.

The problem of discovering information which has a bearing on a given situation has probably received the most attention in recent years, and various mechanical systems have been developed and put into operation. This phase of communication is commonly referred to as *information retrieval* or, more broadly, as the *library problem*. Information retrieval is necessarily a major function of the Business Intelligence System. Means are provided not only to integrate this function with the rest of the system but also to produce additional useful functions, as will be described later.

The achievement of these objectives is governed by principles essential to effective service and convenience of the user. Some of these are listed below:

1. Information admitted to the system includes communications, addressed to action points individually, which contain information of potential interest to other action points.
2. New information which is pertinent or useful to certain action points is selectively disseminated to such points without delay. A function of the system is to present this information to the action point in such a manner

that its existence will be readily recognized.

3. Transmittal of information either as a result of dissemination or of retrieval is to be guided by progressive stages of acceptance by an action point. This procedure saves the recipient's time by reducing the amount of material to be transmitted and eliminating the non-pertinent material.

4. The system is to provide means for quickly discovering similarity of interests and activities that might exist amongst action points so that subjects and problems of common concern may be discussed and advanced through direct interchange of ideas between such points, if so desired.

5. The system is not to impose conditions on its user which require special training to obtain its services. Instead the system is to be operated by experienced library workers. Thus, in the case of an inquiry, the user will be required only to call the librarian, who will accept the query and will ask for any amplification which, in accordance with his experience, will be most helpful in securing the desired information.

6. Similarly, information lingering at an action point but of potential value to other action points is *mobilized* for efficient communication through inquiries of skilled reporters.

#### Description of the Business Intelligence System

The following description is given in rather general terms, and references to any specific type of *business* have been substantially avoided. Furthermore, the fact that certain devices are being referred to as implementation of the system, should not be interpreted as implying a specific size of the operation.

The description is given in accordance with main functional sections of the system, each illustrated by the diagram. Our assembly of these functional sections into a complete system is shown in Fig. 1.

#### Document input

Each document entering the system shown in Fig. 1 is assigned a serial number and is photographically reproduced on some medium such as microfilm. In those cases where the document has been addressed specifically to an action point, the original is promptly transmitted to the addressee. In all other cases the original is stored in a file for a reasonably short time and thereafter destroyed, unless there are reasons for preserving it for longer periods.

The microfilm copy of the document is transcribed onto magnetic tape by a human transcriber or a print-reading device. In those cases where the original document is available in machine-readable form, the transcription is done mechanically. The document is now available both as a microfilm copy and a magnetic tape record.

The microfilm copy is then recopied onto the storage medium of a *document microcopy* storage device. The microfilm record is stored elsewhere to constitute a microfilm master file which may serve to regenerate records in cases of emergency.

The magnetic tape record is now introduced into the auto-abstracting and encoding device. This device submits the document to a statistical analysis based on the physical properties of the text, and data are derived on word frequency and distribution. From these data the device then selects certain sentences of the document to produce an auto-abstract.<sup>2</sup> This is printed out, together with the title, author, and document serial number. This printout is photographically transferred onto the storage medium of the *auto-abstract microcopy storage device*.

The process of creating auto-abstracts consists of ascertaining the frequency of word occurrences in a document. A predetermined portion of the words of highest frequency is then given the status of *significant words* and an analysis is made of all the sentences in the text containing such words. A relative value of sentence significance is then established by a formula which reflects the number of significant words contained in a sentence and the proximity of these words to each other within this sentence. Several sentences which rank highest in value of significance are then extracted from the text to constitute the auto-abstract.

As soon as the auto-abstract has been created, the statistical data are further processed to derive an information pattern which characterizes the document. This process of *encoding* constitutes a further abstraction and involves procedures such as the categorization of words by means of a thesaurus.<sup>3</sup>

Useful patterns may be derived by listing a given portion of the words of highest frequency together with a selection of specific words. The interrelationship of words may also be indicated and certain frequently occurring combinations of words may be noted. Because of variation of word usage amongst authors the normalization of such words becomes an important function of encoding. Index lookup in a thesaurus-like dictionary will replace words, including those of foreign languages, by a notional family designation. The selection of specific words may also be accomplished by index lookup.

The document pattern derived by the above process is then transferred into a special pattern-storage device together with the title, author, and document serial number. This information is stored in coded form on a medium that may be subjected to serial scanning. As an alternative the resulting pattern may be rearranged and be distributed over a storage array to permit random access according to characteristics.

The tape or film transcript of the document may be stored in a library for reference if it later becomes necessary to change the method or scope of encoding.

#### **Action-point profiles**

As indicated earlier, one of the basic requirements of the system is the ability to recognize by mechanical means the sphere of interest and the type of activities that characterize each of the action points the system is to serve. This is accomplished by means of an information pattern similar to that of the documents.

Initially, the creation of these action-point profiles is best accomplished by having each action point create a document describing the various aspects of its activities and enumerating the types of information needed. Such documents are then introduced at the input of the system and are identified by action-point designation. The machine-readable transcripts of these documents are then described in connection with the document input. The resulting patterns are then stored in the Pattern Storage area in a special profile-storage device. Also stored, with each of these profile patterns, is the date of entry.

#### **Selective dissemination of new information**

Based on the document-input operation and the creation of profiles, the system is ready to perform the service function of selective dissemination of new information.

As soon as a new document has been entered into the system and its pattern developed, this pattern is set up in a comparison device which has access to all of the action-point profiles. The comparisons are carried out on the basis of degree of similarity, expressed in terms of a fraction, for each of the profile patterns. This fraction is subject to change as time goes on, depending upon conditions to be explained later.

Whenever a profile agrees to a given extent with a given document pattern, the serial number, title, and author of the affected document, together with the action-point profile designation, are transferred and stored in a monitoring device. This procedure is repeated for any subsequent similar occasion. The monitor is substantially a random-access storage device and has the functional capabilities of performing inventory operations. In this capacity it will transmit the serial number, title and author of the document in question to the desk printer at the selected action point and keep a record of this transaction.

Of the various ways in which such an announcement may be transmitted to the affected action points, the most effective one is by means of a printing device at each action-point location. An objective of the system is to command attention of the recipient. The use of individual printing devices is more effective than are centrally located devices serving several action points.

#### **Selective acceptance of disseminated information**

The dissemination of information so far has consisted in furnishing the action point with the serial number, title, and author of documents selected for it. This selection, however, is considered to be a provisional one, and the system withholds any further information if the action point can determine, on the basis of information given so far, that certain of the selected subjects are not of sufficient interest. If an announcement is of interest, and more detailed information on the subject is desired, the system will produce such information on demand. This step is initiated when the action point connects itself by telephone to the monitor and dials the serial numbers of the documents affected. Upon receipt of this message the monitor will relay an instruction to the microcopy storage

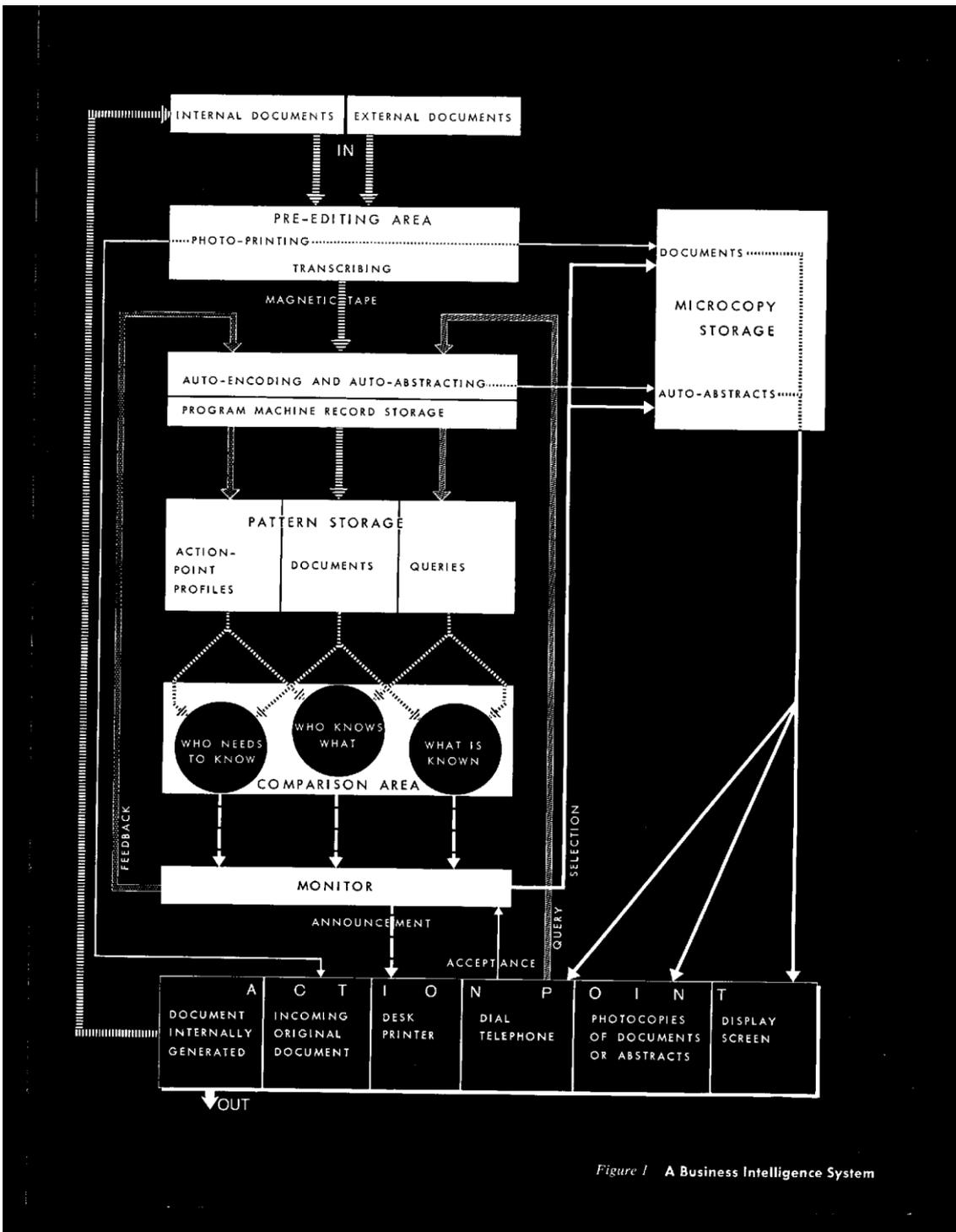


Figure 1 A Business Intelligence System

device to produce photoprints of the auto-abstracts of these documents and to mark them with the action-point designation. The auto-abstracts are then transmitted to the action point either in the form of a paper copy or by speedier means, such as Telefax or TV display.

The action point may now peruse the abstracts to determine which of the documents are desired in their entirety. These decisions are then entered into the system in the form of *acceptances*. An acceptance is made at an action point by dialing the document number, prefixed by a code symbol, whereupon the monitor will instruct the microcopy storage device to produce a photocopy of the complete document, properly marked with the action-point designation. These photocopies are then delivered to the action point.

The monitor will record the incidence of acceptance by modifying the affected records contained in its storage. At the same time the monitor will also instruct the auto-encoding device to transfer copies of the code patterns of the affected documents to the profile section of pattern storage, together with the identification of the action point involved and the date of transferral.

As a result of these operations the profile of a given action point has been updated to reflect interest in a currently communicated subject. As time goes on there is the probability that an increasing number of new documents will be announced to an action point because of possible shift of interests. In order to avoid such cumulative effects, the system is so arranged that the response to past interests is gradually relaxed. This relaxation is related to the date affixed to each new pattern that is superimposed on an action point's profile. Depending on the age of each of these patterns, an adjustment is made on the fraction of similarity that must be met in the comparison process of new documents. The older the profile pattern, the closer an agreement is needed for selection for dissemination, and consequently the fewer documents are selected. On the other hand those documents selected are more closely related to the original subject.

#### **Information retrieval**

This phase of the system concerns itself with the retrieval of those stored documents which might be relevant to a topic under consideration by an action point. The information to be discovered may vary widely and may consist of anything ranging from factual data to an extensive bibliography on a broad subject. Under the supervision of an experienced librarian the process of information retrieval is performed in the following way.

An action point telephones the librarian and states the information wanted. The librarian will then interpret the inquiry and will solicit sufficient background information from the action point in order to provide a document similar in format to that of documents normally entering the system. This query document is transmitted to the auto-encoding device in machine-readable form. An information pattern is then derived from the query document in a manner similar to that used for normal documents.

The resulting query pattern, together with a serial number and designation of the originating action point, is then sent to the queries section of the pattern-storage device. Subsequently, a copy of this query pattern is set up in the comparison device and is compared with all of the document patterns stored in the document-pattern storage device. This operation is similar to the one described in connection with selective dissemination. In the present case, the query pattern replaces the profile pattern.

Whenever similar patterns are detected by this means, the document designation is transmitted to the monitor, where it is registered and then announced to the action point.

Although the service of a librarian is considered a convenience to the action point, in certain cases, means may be provided at the action-point location to permit direct access to the system. This would be justified where many of the inquiries concern lookup-type retrieval of data.

When an action point desires information relative to a given document, the number of the document at hand would be dialed and instructions for search given to the monitor. Thereupon the monitor would select the corresponding pattern from document-pattern storage and provide instruction for use as a query pattern in the ensuing comparison operation.

#### **Selective acceptance of retrieved information**

The considerations which prompted the step-by-step acceptance of documents in the dissemination process are also applied to information retrieval. The processes employed, therefore, are identical.

The function of information retrieval, however, differs from that of dissemination in that the choice is not that of accepting or rejecting one document, but rather a selection of one or several from a special group of potentially relevant documents. Although in some cases a first search may have produced satisfactory references, in other cases the material produced may not be satisfactory. The action point must then relay this fact to the librarian and discuss with him how the searching procedure or the query should be modified so as to improve the probability of getting relevant material.

In those cases where pertinent information has been discovered, the acceptance of the complete documents of such information will cause the updating of the action-point profile, as was the case in dissemination. The query pattern will be impressed on the profile as a matter of course, whether or not the inquiry has been satisfied, so that new documents relevant to the subject of the inquiry will be made known subsequently.

#### **Detection of an action point having given characteristics**

In the process of transacting business it is often desired to determine who concerns himself with a given subject. The usual type of question asked is: "Who does or knows a certain thing?" A function of the Business Intelligence System is to answer questions of this type.

The manner in which this function is performed by the system is similar to the information retrieval procedure. However, instead of simulating a document pattern, a profile pattern is developed which represents most closely the characteristics of an action point sought. This synthetic profile is then compared with those in the profile storage and when a given degree of similarity is discovered, the identification of the affected action point is transferred to the monitor, together with the identification of the inquirer point. Thereafter the identities are announced by the tape-printing device at the inquiring action point so that personal contacts may be made.

#### **Document output**

The functions described so far have concerned themselves with documents admitted or acquired by the system from the outside. The document-output phase deals with internally generated documents. This type of document is essentially the product of action points and may be addressed to other action points within the organization or to external points. An objective of the system is to facilitate selective dissemination and retrieval of such documents in substantially the same way as for outside documents.

When a document has been created at an action point, a copy is produced, preferably in machinable form. This copy is then dispatched for processing to the input point of the system and the original is sent to the addressee.

Since this type of document is an indication of the interest of the originating action point, the information pattern derived by the auto-encoding process is not only stored in document-pattern storage but also is impressed on the profile of its originator, thereby updating it.

In the dissemination process this internally created document is announced to other action points in the same fashion as were outside documents.

#### **Miscellaneous functions of the system**

The comprehensive system for the various functions so far described is illustrated by Fig. 1. A number of additional useful functions which may be derived from the system are briefly described here.

It might be desirable to check each new document for duplication by comparing it with all of the documents in storage. Similarly a list of related documents may be prepared to serve as references applying to a new document.

When retrieving information it might be found advantageous to compare a query first with all the queries stored, in order to discover whether similar queries have been submitted in the past. If a list of the documents retrieved is available, the process of retrieval may be greatly simplified. This method may also be used to bring together the respective inquirers to furnish an opportunity to discuss the problems which apparently brought about similar inquiries. Periodic analysis of the profiles may also furnish valuable information on trends and possible overlapping of activities or interests.

Since a history of the usage of the system is stored in the monitor, an analysis of its records will disclose the efficiency of system operation. The findings may serve to adjust the system for optimum efficiency.

There are many details which might have to be provided to adjust the general form of the system to specific applications. One such requirement might be classification, by an editor, of documents with regard to security, proprietary interests and proper utilization of information.

A plurality of systems may be organized in hierarchical fashion, in which a first system would serve a number of more specialized systems. In this case the specialized system would each assume the role of an action point in the mother system.

It also appears quite feasible to share the system equipment among a number of organizations.

#### **Prospects for establishing a Business Intelligence System**

The system described here employs rather advanced design techniques and the question arises as to how far away such systems may be from realization. It may therefore be of interest to review the state of system and machine development.

The availability of documents in machine-readable form is a basic requirement of the system. Typewriters with paper-tape punching attachments are already used extensively in information processing and communication operations. Their use as standard equipment in the future would provide machine-readable records of new information. The transcription of old records would pose a problem, since in most cases it would be uneconomical to perform this job by hand. The mechanization of this operation will therefore have to wait until print-reading devices have been perfected.

The type of equipment required for processing information in accordance with the system is presently available as far as the functions are concerned. It is safe to assume that special equipment will eventually be required to optimize the operation.

The auto-abstracting and auto-encoding systems are in their early stage of development and a great deal of research has yet to be done to perfect them. Perhaps the techniques which ultimately find greatest use will bear little resemblance to those now visualized, but some form of automation will ultimately provide an effective answer to business intelligence problems.

#### **References**

1. *Webster's New Collegiate Dictionary*, G. & C. Merriam Co., Springfield, Mass.
2. H. P. Luhn, "The Automatic Creation of Literature Abstracts," *IBM Journal of Research and Development*, **2**, No. 2, 159 (April 1958).
3. H. P. Luhn, "A Statistical Approach to Mechanized Encoding and Searching of Literary Information," *IBM Journal of Research and Development*, **1**, No. 4, 309 (October 1957).

Received July 1, 1958

319

# Apêndice 2

Artigo – John McCarthy

A PROPOSAL FOR THE  
DARTMOUTH SUMMER RESEARCH PROJECT  
ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College  
M. L. Minsky, Harvard University  
N. Rochester, I.B.M. Corporation  
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

The following are some aspects of the artificial intelligence problem:

### **1 Automatic Computers**

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

### **2. How Can a Computer be Programmed to Use a Language**

It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning and rules of conjecture. From this point of view, forming a generalization consists of admitting a new word and some rules whereby sentences containing it imply and are implied by others. This idea has never been very precisely formulated nor have examples been worked out.

### **3. Neuron Nets**

How can a set of (hypothetical) neurons be arranged so as to form concepts. Considerable theoretical and experimental work has been done on this problem by Uttley, Rashevsky and his group, Farley and Clark, Pitts and McCulloch, Minsky, Rochester and Holland, and others. Partial results have been obtained but the problem needs more theoretical work.

### **4. Theory of the Size of a Calculation**

If we are given a well-defined problem (one for which it is possible to test mechanically whether or not a proposed answer is a valid answer) one way of solving it is to try all possible answers in order. This method is inefficient, and to exclude it one must have some criterion for efficiency of calculation. Some consideration will show that to get a measure of the efficiency of a calculation it is necessary to have on hand a method of measuring the complexity of calculating devices which in turn can be done if one has a theory of the complexity of functions. Some partial results on this problem have been obtained by Shannon, and also by McCarthy.

### **5. Self-Improvement**

Probably a truly intelligent machine will carry out activities which may best be described as self-improvement. Some schemes for doing this have been proposed and are worth further study. It seems likely that this question can be studied abstractly as well.

### **6. Abstractions**

A number of types of "abstraction" can be distinctly defined and several others less distinctly. A direct attempt to classify these and to describe machine methods of forming abstractions from sensory and other data would seem worthwhile.

### **7. Randomness and Creativity**

A fairly attractive and yet clearly incomplete conjecture is that the difference between creative thinking and unimaginative competent thinking lies in the injection of a some randomness. The randomness must be guided by intuition to be efficient. In other words, the educated guess or the hunch include controlled randomness in otherwise orderly thinking.

In addition to the above collectively formulated problems for study, we have asked the individuals taking part to describe what they will work on. Statements by the four originators of the project are attached.

We propose to organize the work of the group as follows.

Potential participants will be sent copies of this proposal and asked if they would like to work on the artificial intelligence problem in the group and if so what they would like to work on. The invitations will be made by the organizing committee on the basis of its estimate of the individual's potential contribution to the work of the group. The members will circulate their previous work and their ideas for the problems to be attacked during the months preceding the working period of the group.

During the meeting there will be regular research seminars and opportunity for the members to work individually and in informal small groups.

The originators of this proposal are:

1. **C. E. Shannon**, Mathematician, Bell Telephone Laboratories. Shannon developed the statistical theory of information, the application of propositional calculus to switching circuits, and has results on the efficient synthesis of switching circuits, the design of machines that learn, cryptography, and the theory of Turing machines. He and J. McCarthy are co-editing an Annals of Mathematics Study on "The Theory of Automata".

2. **M. L. Minsky**, Harvard Junior Fellow in Mathematics and Neurology. Minsky has built a machine for simulating learning by nerve nets and has written a Princeton PhD thesis in mathematics entitled, "Neural Nets and the Brain Model Problem" which includes results in learning theory and the theory of random neural nets.

3. **N. Rochester**, Manager of Information Research, IBM Corporation, Poughkeepsie, New York. Rochester was concerned with the development of radar for seven years and computing machinery for seven years. He and another engineer were jointly responsible for the design of the IBM Type 701 which is a large scale automatic computer in wide use today. He worked out some of the automatic programming techniques which are in wide use today and has been concerned with problems of how to get machines to do tasks which previously could be done only by people. He has also worked on simulation of nerve nets with particular emphasis on using computers to test theories in neurophysiology,

4. **J. McCarthy**, Assistant Professor of Mathematics, Dartmouth College. McCarthy has worked on a number of questions connected with the mathematical nature of the thought process including the theory of Turing machines, the speed of computers, the relation of a brain model to its environment, and the use of languages by machines. Some results of this work are included in the forthcoming "Annals Study" edited by Shannon and McCarthy. McCarthy's other work has been in the field of differential equations.

The Rockefeller Foundation is being asked to provide financial support for the project on the following basis:

1. Salaries of \$1200 for each faculty level participant who is not being supported by his own organization. It is expected, for example, that the participants from Bell Laboratories and IBM Corporation will be supported by these organizations while those from Dartmouth and Harvard will require foundation support.

2. Salaries of \$700 for up to two graduate students.

3. Railway fare for participants coming from a distance.

4. Rent for people who are simultaneously renting elsewhere.

5. Secretarial expenses of \$650, \$500 for a secretary and \$150 for duplicating expenses.

6. Organization expenses of \$200. (Includes expense of reproducing preliminary work by participants and travel necessary for organization purposes.

7. Expenses for two or three people visiting for a short time.

Estimated Expenses

6 salaries of 1200	\$7200
2 salaries of 700	1400
8 traveling and rent expenses averaging 300	2400
Secretarial and organizational expense	850
Additional traveling expenses	600
Contingencies	550
	<hr/>
	\$13,500

## PROPOSAL FOR RESEARCH BY C. E. SHANNON

I would like to devote my research to one or both of the topics listed below. While I hope to do so, it is possible that because of personal considerations I may not be able to attend for the entire two months. I, nevertheless, intend to be there for whatever time is possible,

1. Application of information theory concepts to computing machines and brain models. A basic problem in information theory is that of transmitting information reliably over a noisy channel. An analogous problem in computing machines is that of reliable computing using unreliable elements. This problem has been studied by von Neumann for Sheffer stroke elements and by Shannon and Moore for relays; but there are still many open questions. The problem for several elements, the development of concepts similar to channel capacity, the sharper analysis of upper and lower bounds on the required redundancy, etc. are among the important issues. Another question deals with the theory of information networks where information flows in many closed loops (as contrasted with the simple one-way channel usually considered in communication theory). Questions of delay become very important in the closed loop case, and a whole new approach seems necessary. This would probably involve concepts such as partial entropies when a part of the past history of a message ensemble is known.

2. The matched environment - brain model approach to automata. In general a machine or animal can only adapt to or operate in a limited class of environments. Even the complex human brain first adapts to the simpler aspects of its environment, and gradually builds up to the more complex features. I propose to study the synthesis of brain models by the parallel development of a series of matched (theoretical) environments and corresponding brain models which adapt to them. The emphasis here is on clarifying the environmental model, and representing it as a mathematical structure. Often in discussing mechanized intelligence, we think of machines performing the most advanced human thought activities—proving theorems, writing music, or playing chess. I am proposing here to start at the simple and when the environment is neither hostile (merely indifferent) nor complex, and to work up through a series of easy stages in the direction of these advanced activities.

## PROPOSAL FOR RESEARCH BY M. L. MINSKY

It is not difficult to design a machine which exhibits the following type of learning. The machine is provided with input and output channels and an internal means of providing varied output responses to inputs in such a way that the machine may be "trained" by a "trial and error" process to acquire one of a range of input-output functions. Such a machine, when placed in an appropriate environment and given a criterion of "success" or "failure" can be trained to exhibit "goal-seeking" behavior. Unless the machine is provided with, or is able to develop, a way of abstracting sensory material, it can progress through a complicated environment only through painfully slow steps, and in general will not reach a high level of behavior.

Now let the criterion of success be not merely the appearance of a desired activity pattern at the output channel of the machine, but rather the performance of a given manipulation in a given environment. Then in certain ways the motor situation appears to be a dual of the sensory situation, and progress can be reasonably fast only if the machine is equally capable of assembling an ensemble of "motor abstractions" relating its output activity to changes in the environment. Such "motor abstractions" can be valuable only if they relate to changes in the environment which can be detected by the machine as changes in the sensory situation, i.e., if they are related, through the structure of the environment, to the sensory abstractions that the machine is using.

I have been studying such systems for some time and feel that if a machine can be designed in which the sensory and motor abstractions, as they are formed, can be made to satisfy certain relations, a high order of behavior may result. These relations involve pairing, motor abstractions with sensory abstractions in such a way as to produce new sensory situations representing the changes in the environment that might be expected if the corresponding motor act actually took place.

The important result that would be looked for would be that the machine would tend to build up within itself an abstract model of the environment in which it is placed. If it were given a problem, it could first explore solutions within the internal abstract model of the environment and then attempt external experiments. Because of this preliminary internal study, these external experiments would appear to be rather clever, and the behavior would have to be regarded as rather "imaginative"

A very tentative proposal of how this might be done is described in my dissertation and I intend to do further work in this direction. I hope that by summer 1956 I will have a model of such a machine fairly close to the stage of programming in a computer.

## PROPOSAL FOR RESEARCH BY N. ROCHESTER

### Originality in Machine Performance

In writing a program for an automatic calculator, one ordinarily provides the machine with a set of rules to cover each contingency which may arise and confront the machine. One expects the machine to follow this set of rules slavishly and to exhibit no originality or common sense. Furthermore one is annoyed only at himself when the machine gets confused because the rules he has provided for the machine are slightly contradictory. Finally, in writing programs for machines, one sometimes must go at problems in a very laborious manner whereas, if the machine had just a little intuition or could make reasonable guesses, the solution of the problem could be quite direct. This paper describes a conjecture as to how to make a machine behave in a somewhat more sophisticated manner in the general area suggested above. The paper discusses a problem on which I have been working sporadically for about five years and which I wish to pursue further in the Artificial Intelligence Project next summer.

### The Process of Invention or Discovery

Living in the environment of our culture provides us with procedures for solving many problems. Just how these procedures work is not yet clear but I shall discuss this aspect of the problem in terms of a model suggested by Craik<sup>1</sup>. He suggests that mental action consists basically of constructing little engines inside the brain which can simulate and thus predict abstractions relating to environment. Thus the solution of a problem which one already understands is done as follows:

1. The environment provides data from which certain abstractions are formed.
2. The abstractions together with certain internal habits or drives provide:
  - 2.1 A definition of a problem in terms of desired condition to be achieved in the future, a goal.
  - 2.2 A suggested action to solve the problem.
  - 2.3 Stimulation to arouse in the brain the engine which corresponds to this situation.
3. Then the engine operates to predict what this environmental situation and the proposed reaction will lead to.
4. If the prediction corresponds to the goal the individual proceeds to act as indicated.

The prediction will correspond to the goal if living in the environment of his culture has provided the individual with the solution to the problem. Regarding the individual as a stored program calculator, the program contains rules to cover this particular contingency.

For a more complex situation the rules might be more complicated. The rules might call for testing each of a set of possible actions to determine which provided the solution. A still more complex set of rules might provide for uncertainty about the environment, as for example in playing tic tac toe one must not only consider his next move but the various possible moves of the environment (his opponent).

Now consider a problem for which no individual in the culture has a solution and which has resisted efforts at solution. This might be a typical current unsolved scientific problem. The individual might try to solve it and find that every reasonable action led to failure. In other words the stored program contains rules for the solution of this problem but the rules are slightly wrong.

In order to solve this problem the individual will have to do something which is unreasonable or unexpected as judged by the heritage of wisdom accumulated by the culture. He could get such behavior by trying different things at random but such an approach would usually be too inefficient. There are usually too many possible courses of action of which only a tiny fraction are acceptable. The individual needs a hunch, something unexpected but not altogether reasonable. Some problems, often those which are fairly new and have not resisted much effort, need just a little

randomness. Others, often those which have long resisted solution, need a really bizarre deviation from traditional methods. A problem whose solution requires originality could yield to a method of solution which involved randomness.

In terms of Craik's<sup>1</sup>S model, the engine which should simulate the environment at first fails to simulate correctly. Therefore, it is necessary to try various modifications of the engine until one is found that makes it do what is needed.

Instead of describing the problem in terms of an individual in his culture it could have been described in terms of the learning of an immature individual. When the individual is presented with a problem outside the scope of his experience he must surmount it in a similar manner.

So far the nearest practical approach using this method in machine solution of problems is an extension of the Monte Carlo method. In the usual problem which is appropriate for Monte Carlo there is a situation which is grossly misunderstood and which has too many possible factors and one is unable to decide which factors to ignore in working out analytical solution. So the mathematician has the machine making a few thousand random experiments. The results of these experiments provide a rough guess as to what the answer may be. The extension of the Monte Carlo Method is to use these results as a guide to determine what to neglect in order to simplify the problem enough to obtain an approximate analytical solution.

It might be asked why the method should include randomness. Why shouldn't the method be to try each possibility in the order of the probability that the present state of knowledge would predict for its success? For the scientist surrounded by the environment provided by his culture, it may be that one scientist alone would be unlikely to solve the problem in his life so the efforts of many are needed. If they use randomness they could all work at once on it without complete duplication of effort. If they used system they would require impossibly detailed communication. For the individual maturing in competition with other individuals the requirements of mixed strategy (using game theory terminology) favor randomness. For the machine, randomness will probably be needed to overcome the shortsightedness and prejudices of the programmer. While the necessity for randomness has clearly not been proven, there is much evidence in its favor.

#### The Machine With Randomness

In order to write a program to make an automatic calculator use originality it will not do to introduce randomness without using foresight. If, for example, one wrote a program so that once in every 10,000 steps the calculator generated a random number and executed it as an instruction the result would probably be chaos. Then after a certain amount of chaos the machine would probably try something forbidden or execute a stop instruction and the experiment would be over.

Two approaches, however, appear to be reasonable. One of these is to find how the brain manages to do this sort of thing and copy it. The other is to take some class of real problems which require originality in their solution and attempt to find a way to write a program to solve them on an automatic calculator. Either of these approaches would probably eventually succeed. However, it is not clear which would be quicker nor how many years or generations it would take. Most of my effort along these lines has so far been on the former approach because I felt that it would be best to master all relevant scientific knowledge in order to work on such a hard problem, and I already was quite aware of the current state of calculators and the art of programming them.

The control mechanism of the brain is clearly very different from the control mechanism in today's calculators. One symptom of the difference is the manner of failure. A failure of a calculator characteristically produces something quite unreasonable. An error in memory or in data transmission is as likely to be in the most significant digit as in the least. An error in control can do nearly anything. It might execute the wrong instruction or operate a wrong input-output unit. On the other hand human errors in speech are apt to result in statements which almost make sense (consider someone who is almost asleep, slightly drunk, or slightly feverish). Perhaps the

mechanism of the brain is such that a slight error in reasoning introduces randomness in just the right way. Perhaps the mechanism that controls serial order in behavior<sup>2</sup> guides the random factor so as to improve the efficiency of imaginative processes over pure randomness.

Some work has been done on simulating neuron nets on our automatic calculator. One purpose was to see if it would be thereby possible to introduce randomness in an appropriate fashion. It seems to have turned out that there are too many unknown links between the activity of neurons and problem solving for this approach to work quite yet. The results have cast some light on the behavior of nets and neurons, but have not yielded a way to solve problems requiring originality.

An important aspect of this work has been an effort to make the machine form and manipulate concepts, abstractions, generalizations, and names. An attempt was made to test a theory<sup>3</sup> of how the brain does it. The first set of experiments occasioned a revision of certain details of the theory. The second set of experiments is now in progress. By next summer this work will be finished and a final report will have been written.

My program is to try next to write a program to solve problems which are members of some limited class of problems that require originality in their solution. It is too early to predict just what stage I will be in next summer, or just; how I will then define the immediate problem. However, the underlying problem which is described in this paper is what I intend to pursue. In a single sentence the problem is: how can I make a machine which will exhibit originality in its solution of problems?

#### REFERENCES

1. K.J.W. Craik, The Nature of Explanation, Cambridge University Press, 1943 (reprinted 1952), p. 92.
2. K.S. Lashley, "The Problem of Serial Order in Behavior", in Cerebral Mechanism in Behavior, the Hixon Symposium, edited by L.A. Jeffress, John Wiley & Sons, New York, pp. 112-146, 1951.
3. D. O. Hebb, The Organization of Behavior, John Wiley & Sons, New York, 1949