

TC 010.259/2018-0

Interessados e matrículas:

Aloísio Dourado Neto (8814-5)

Giuseppe de Abreu Antonaci (10.695-0)

Luiz Henrique Batistuta Gomide (9981-3)

Rodrigo Otávio Coelho Hildebrand
(9466-8)

Assunto:

Forum Global SAS 2018 em Dever,
Estados Unidos, no período de 8 a 11 de
abril de 2018.

RELATÓRIO CIRCUNSTANCIADO

INTRODUÇÃO

1. Trata-se de relatório circunstanciado referente ao Forum Global SAS 2018 em Denver, Estados Unidos, no período de 8 a 11 de abril de 2018.
 2. O tema de Fiscalização Contínua de Benefícios Sociais (FCB) faz parte da estratégia específica da Situação-problema: 10. Deficientes mecanismos estruturais para o combate à corrupção na administração pública, que foi escolhida pelo Presidente como uma das situações problemas que serão priorizadas pelo Tribunal.
 3. A SecexPrevidência, desde de 2015, realiza Fiscalização Contínua na concessão, manutenção e pagamento dos benefícios previdenciários, trabalhistas e assistências, que juntos superam o volume de pagamentos de R\$ 650 bilhões ao ano. A FCB depende, dentre outras ferramentas, do uso intensivo do módulo de detecção de fraudes da empresa de software SAS, que atualmente é utilizada apenas pela SecexPrevidência, esses ciclos da FCB já possibilitaram mais de R\$ 2 bilhões de benefícios potenciais de controle.
 4. Dessa forma, como os quatro auditores envolvidos realizam análise de qualidade, cruzamentos de bases de dados, análise da credibilidade, criação de modelos com machine learning para benefícios diferentes (trabalhistas, previdenciários e assistenciais), com tipologias e complexidades díspares, faz-se necessário que todos participem do referido evento da SAS, haja vista que os interesses de cada área são diferentes e que ocorrerão palestras em momentos simultâneos e que os auditores precisarão se organizar para poderem atender a essas palestras.
 5. Dessa forma, entende-se ser garantida a absorção de informações, de conhecimentos e de desenvolvimento de competências dos auditores da SecexPrevidência participantes (Aloísio para o seguro desemprego e abono salarial; Rodrigo aposentadorias, auxílios e pensões; Giuseppe Programa Bolsa Família; Luiz Henrique Benefício de Prestação Continuada e Cadastro Único).
 6. O evento, promovido pela desenvolvedora do SAS, um dos mais importantes softwares de análise de dados, tem como objetivo compartilhar conhecimentos e experiência entre os vários analistas de dados do mundo, promovendo o uso cada vez mais eficiente do SAS em busca de resultados melhores e de maior utilidade. Além das palestras que servirão de inspiração para as análises da FCB, a apresentação a ser realizada pela equipe demonstra o nível de excelência e o protagonismo que queremos liderar no Brasil no que tange a controle de benefícios.
-

7. A participação no evento ajudará que os servidores, que já executam análises de dados nas áreas de Previdência Social, Assistência Social e Trabalho com o SAS, tenham a oportunidade de trocar experiências e adquirir conhecimentos com outros analistas, fomentando novas possibilidades de uso da ferramenta e a obtenção de mais resultados no âmbito do controle externo e da Fiscalização Contínua de Benefícios.

8. As técnicas aprendidas durante o evento, bem como, as experiências de implementação de sistemas de pagamentos indevidos, serão de grande valia nas FCB baseadas em análises de dados executadas pelos servidores na SecexPrevidência. Assim, entende-se que os servidores adquirirão maior conhecimento sobre possibilidades de uso da ferramenta, com o objetivo de aprimorar e otimizar o trabalho já realizado.

9. Os servidores atuam no Núcleo de Análise de Dados e Tecnologia da Informação (Ncad) da SecexPrevidência e vêm realizando fiscalizações baseadas em análise de dados. Dessa maneira, os componentes do núcleo terão a oportunidade de compartilhar os conhecimentos in loco, na implementação de novas metodologias de fiscalização.

10. Ademais, encontra-se em elaboração o Sistema de Controle de Pagamentos (módulo benefícios) no qual se buscará disseminar a tecnologia desenvolvida e implementada pela Secex Previdência nas FC para outras áreas do Tribunal. Para tanto, a equipe que participará do evento tem a incumbência de conceber e apoiar todo o desenvolvimento do referido Sistema que significará uma revolução para o Tribunal, que passará a atuar de forma contínua na fiscalização de todos os benefícios pagos pela administração pública federal.

11. Em suma, entende-se que a participação dos quatro servidores no referido evento é fundamental para a implementação de novas tipologias e na criação do Sistema de Controle de Pagamentos (módulo benefícios), representando um investimento crucial para garantir o envolvimento e o conhecimento técnico necessários para viabilizar a continuação e a expansão da fiscalização contínua no âmbito do tribunal.

RELATO

12. No período de 08 a 11 de abril deste ano realizou-se na cidade de Denver, Colorado, evento denominado *SAS Global Forum 2018*. O evento, promovido pela desenvolvedora do SAS, um dos mais importantes softwares de análise de dados, tem como objetivo compartilhar conhecimentos e experiência entre os vários analistas de dados do mundo, promovendo o uso cada vez mais eficiente do SAS em busca de resultados melhores e de maior utilidade. Neste ano de 2018, o evento contou com aproximadamente 6 mil inscritos, 612 apresentações, 94 pôsteres e 20 sessões de workshop. A participação brasileira incluiu empresas privadas de setores de tecnologia, varejo, bancos e seguradoras, assim como órgãos e entidades públicas como Ministério do Planejamento, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), Universidade de Brasília (UnB), Tribunal de Contas do Distrito Federal (TCDF) e Tribunal de Contas do Estado do Ceará (TCE-CE).

13. Pelo TCU, participaram os auditores Aloísio Dourado, Giuseppe Antonaci, Luiz Henrique Gomide e Rodrigo Hildebrand do Núcleo de Análise de Dados e Tecnologia da Informação (Ncad) da SecexPrevidência. A secretaria, desde de 2015, realiza Fiscalização Contínua de Benefícios (FCB) na concessão, manutenção e pagamento dos benefícios previdenciários, trabalhistas e assistências, que juntos superam o volume de pagamentos de R\$ 650 bilhões ao ano, correspondendo a cerca de 66% do orçamento federal, excluídos os Encargos Especiais. A FCB depende, dentre outras ferramentas, do uso intensivo do módulo de detecção de fraudes da empresa de software SAS, que atualmente é utilizada apenas pela SecexPrevidência. Esses ciclos da FCB já possibilitaram mais de R\$ 2 bilhões em benefícios potenciais e efetivos de controle.

14. Dentre os diversos tópicos relacionados a Tecnologia de Informação apresentados no evento, citam-se aqueles mais afetos ao trabalho realizado pelo Ncad, como qualidade, integração, mineração e visualização de dados, modelagem preditiva, análise de texto e integração de software livre.

15. Como a dinâmica do evento previa a realização de diversas sessões paralelas, cada participante se responsabilizou por organizar sua própria agenda, mas buscando cobrir a maior gama possível de assuntos simultâneos. Durante as discussões das agendas individuais, cobriram-se apresentações, palestras e discussões de dois programas relacionados, apesar de paralelos: Programa de Usuários e Programa Executivo.

16. Na sequência do relatório é apresentada uma breve descrição das apresentações julgadas mais pertinentes ao trabalho realizado pelo Ncad e também aquelas que podem trazer ganhos a outras áreas do TCU. Ressalta-se que algumas das apresentações relatadas usam análise de dados para combate à fraude ou avaliação de políticas públicas em outros assuntos diferentes da Previdência, como Educação e Saúde, sendo pertinentes às respectivas áreas.

Estrutura do evento

17. O *SAS Global Forum 2018* iniciou-se no dia 8 de abril, com Sessão Geral de abertura, em que o CEO da empresa, Dr. Jim Goodnight, e outros executivos trouxeram novidades e detalhes do evento.

18. Nos três dias seguintes, houve a realização paralela dos dois programas, cada um com sua agenda específica, que continham Sessões Gerais (apresentações que geralmente não concorriam com outras sessões); *Breakout Sessions*, apresentações de pequena duração para grupos de poucas pessoas (muitas apresentações ocorriam paralelamente); o “*The Quad*”, local em que se disponibilizaram *e-posters*, demonstrações, espaços para *networking* e apresentações de produtos; e treinamentos.

Descrição das Sessões Gerais

19. As Sessões Gerais trouxeram nomes de prestígio em diversas áreas para compartilhar suas experiências nas áreas de Análise de Dados, Liderança ou Negócios.

20. A primeira Sessão Geral, a de abertura do evento, trouxe diversos casos de sucesso do uso de Análise de Dados em diferentes áreas de negócio. Relataram-se iniciativas em análise de venda (preço) e localização de bilhetes (em estádios) para eventos esportivos; reconhecimento de imagem para detecção de parâmetros relacionados a jogadas de futebol; análise de imagens médicas (avaliação de fluxo sanguíneo em três dimensões); aspectos de preparação de dados e sigilo; análises de *cluster* e processos de inteligência; e acessibilidade a dados “sonográficos” pela Apple e SAS (mecanismos para ampliar a acessibilidade em Análise de Dados).

21. Em seguida, Dan Heath, autor *best-seller* do New York Times, trouxe algumas das ideias de seu livro *The Power of Moments*, explorando por que e como certas experiências se destacam das demais e nos mudam. Ele trouxe os quatro elementos que definem tais momentos, na vida pessoal e profissional: *elevation*, *insight*, *pride* e *connection*, trazendo diversos exemplos (parques da Disney e o segundo hotel mais bem avaliado pelo Tripadvisor na Califórnia).

22. Outro *best-seller* a apresentar uma Sessão Geral foi o ganhador do Pulitzer Jon Meacham, escritor de *The American Odyssey of George Herbert Walker Bush*. Em sua apresentação, *The Art of Leadership*, o escritor destaca as principais características de liderança, como a curiosidade e a empatia, fazendo relação direta ao trabalho com Análise de Dados e ao ambiente empresarial, como um todo.

23. Ainda no tema de Liderança, Carey Lohrenz trouxe a apresentação *Lessons in Leadership*, em que abordou sua trajetória como primeira piloto mulher de um F-14 *fighter* na Marinha americana, escritora e CEO. Entre outros pontos, destacou trabalhar em ambientes extremos e de alta pressão, colocando o papel essencial que as equipes têm em uma organização. Como exemplo pessoal, relatou a experiência de trabalhar em um porta-aviões com equipes de jovens (idade média de dezenove anos) e com pouco tempo de entrosamento (alta rotatividade).

24. Por fim, Peyton Manning, um atleta da NFL com recordes em *touchdowns* e *passing yards*, além de dois *Superbowls* (e participação pelo time regional), participou de uma mesa com Randy Guard, Vice-presidente executivo e *Chief Marketing Officer* da SAS, onde abordaram vários eventos da carreira do jogador, trazendo inspirações em liderança, motivação, trabalho em equipe e processo de tomada de decisão.

Descrição das *Breakout Sessions* e *e-posters*

25. A breve descrição das sessões e dos pôsteres usou como base também os artigos disponibilizados no site do fórum¹.

TCU: Breakout Session 2573-2018. SAS Fraud Framework and MCMC in Government Estimation of Improper Payment of Social Benefits

26. Em uma *Breakout Session*, a Secex Previdência apresentou uma iniciativa desenvolvida em 2017 com o uso do SAS Fraud Framework e da técnica estatística de simulações de Monte Carlo via cadeias de Markov para avaliar a percepção de pagamentos indevidos realizados pelo INSS. Por meio de entrevistas a especialistas, a equipe coletou percepções individuais de padrões de fraude para cada grupo de benefícios pago pelo INSS (além de benefícios tipicamente previdenciários, como aposentadorias por idade, por tempo de contribuição e pensões por morte, incluíram-se os benefícios assistenciais BPC Idoso e BPC deficiente físico). Após as etapas de Preparação dos dados, Análise e Simulação, estimou-se o índice de pagamentos indevidos como sendo, no mínimo, de 11,41% dos benefícios pagos pelo INSS. Esse trabalho insere-se no acompanhamento de Fiscalização Contínua.

Maxtera: Breakout Session 1729-2018. Results of the Application of SAS for Fraud Identification in the Public Administration of Brazil

27. Os autores apresentaram os trabalhos realizados pela empresa Maxtera em relação aos clientes governamentais. Foram apresentados os resultados dos trabalhos de fraude no sistema de Previdência realizado no TCU, na identificação de fraude e sonegação de impostos na Secretaria de Fazenda do Mato Grosso, os cruzamentos com folha de pagamento no Tribunal de Contas do Distrito Federal (TCDF) e do Tribunal de Contas do Ceará.

TCDF: E-poster 2723-2018. Using SAS Fraud Framework for Government to Identify Fraud in Brazil's Federal Capital

28. Os autores apresentaram os resultados dos cruzamentos de dados realizados com as informações de servidores do Distrito Federal para evitar a acumulação indevida e os relacionamentos entre empresas, pessoas e funcionários distritais para as licitações e os contratos do governo distrital.

TCE-Ceará: Breakout Session 2596-2018. Indication of Irregularities in the Accumulation of Public Jobs in the State of Ceará (Brazil)

¹ <http://support.sas.com/resources/papers/proceedings18/>

29. O Tribunal de Contas do Estado do Ceará apresentou um trabalho em que buscou identificar servidores que possuíam acumulação irregular de emprego público ou acumulação com empregos privados que ultrapassavam 60 horas semanais de carga horária. A partir dessa tipologia o TCE-Ceará está conseguindo evitar o pagamento indevido de aproximadamente 8 milhões de dólares.

MPOG: Breakout Session 2707-2018. Identifying Behaviours and Irregularities in Public Procurement in the Brazilian Federal Government

30. O Ministério do Planejamento, Desenvolvimento e Gestão apresentou um estudo realizado nas licitações eletrônicas executadas pela administração pública. Uma das situações de fraude que o estudo buscou identificar foi empresas que declararam ser micro ou pequena empresa com o objetivo de obter vantagem no certame quando não possuíam tal condição. Outra situação avaliada foi o uso de robôs para realizar lances nos leilões eletrônicos, o que pode levar a uma vantagem indevida da empresa operando esses robôs, ainda que não se tenha avaliado se há desvantagem para a administração nesse caso. Por fim, foi avaliado se há evidências de conluio entre as empresas concorrentes que participam do mesmo leilão a partir da análise dos seus comportamentos nos diversos certames.

Breakout Session 2052-2018. Enable Personal Data Governance for Sustainable Compliance

31. Os autores apresentaram a Regulamentação da União Europeia para Proteção aos Dados, *European Union's General Data Protection Regulation* (EU GDPR), e quais os principais desafios dos controladores de dados e os processadores de dados para demonstrar o *compliance* das obrigações estabelecidas pela regulamentação, por meio da documentação dos processos e do levantamento dos riscos desses processos para os indivíduos.

32. Foram apresentados o estado atual do sistema em desenvolvimento para mapear as ações de *compliance* e a verificação das categorias de dados pessoais que são armazenados.

Breakout Session 2396-2018. Using Arrays to Quickly Perform Fuzzy Merge Look-Ups: Case Studies in Efficiency

33. O termo *Fuzzy Merge* se refere ao processo de cruzar duas bases de dados quando não existe um campo chave que relacione ambas as bases. Os autores apresentam formas de realizar um *Fuzzy Merge* utilizando uma matriz para melhorar a eficiência no processamento dos comandos.

Breakout Session 1746. Merge with Caution: How to Avoid Common Problems When Combining SAS Data Sets

34. Foram apresentadas na sessão as principais causas para erros no cruzamento de entre bases de dados utilizando o SAS. As principais causas são: i) esquecer a declaração BY; ii) usar a declaração SET em vez de MERGE; iii) usar a declaração BY para colunas com informações diferentes; iv) cruzamento de N para N; v) Campos usados na declaração BY com tamanhos diferentes; vi) variáveis com sobreposição.

Breakout Session 1684. Adopt A Pet (Elephant?): Are You Enjoying Your Apache Hadoop Investment?

35. O autor realizou uma revisão sobre o sistema Hadoop, sua arquitetura e as considerações que devem ser realizadas sobre o *return of investment* ao adotar o sistema para

a armazenagem dos dados. Foram mostradas as interações entre o ambiente Hadoop e o software SAS, bem como exemplos de código para realizar procedimentos.

Breakout Session 2160. Working with Big Data in SAS

36. A apresentação mostrou os desafios e as soluções para utilizar o software SAS com grandes bases de dados. A partir da relação base entre o software SAS, o acesso a base de dados no disco e o processamento dos comandos, foram mostrados a evolução e formas de otimizar esse relacionamento. Foram mostrados exemplos que o processamento em SQL por meio de *pass-through* melhora a performance; na versão 9.4 do software SAS, o processamento em DS2 permite o processamento em paralelo e a capacidade de processamento e acesso ao disco em paralelo com a utilização do *Cloud Analytic Service* da nova versão do software SAS.

Session 2693. Enterprise, Prepare for Analysis! Using SAS Data Management to Prepare Data for Analysis

37. Foi apresentado um demo do software *SAS Data Management* nos processos de integração de bases de dados de diferentes fontes, limpeza das informações e enriquecimento da base para análises.

Breakout Session 2886. Fuzzy Matching Programming Techniques Using SAS Software

38. O termo *Fuzzy Merge* se refere ao processo de cruzar duas bases de dados quando não existe um campo chave que relacione ambas as bases. Os autores apresentaram o conceito e *fuzzy matching* e a utilização das funções de SOUNDEX (transformação de texto em fonemas), COMPLEV (distância de levenshtein), SPEDIS e COMPGED (funções de distâncias de *strings*), para realizar o cruzamento de informações em bancos de dados cujos campos chaves não são comuns entre as bases.

Breakout Session 1884. Comparing SAS Viya and SAS 9.4 Capabilities: A Tale of Two SAS Platform Engines

39. A empresa SAS está desenvolvendo uma nova versão do software SAS, chamado de SAS Viya. Nessa sessão, foram apresentadas as diferenças entre as versões SAS 9.4 e SAS Viya. Foram mostradas as diferenças na arquitetura e na capacidade de processamento em nuvem do software SAS Viya e as modificações da nova versão que foram portadas para a versão anterior, como o aumento dos campos de nome das variáveis e a funcionalidade de utilizar a linguagem python junto com o SAS.

Hands on Workshop: Macro coding by example

40. O workshop sobre programação em linguagem SAS, mostrou os principais erros cometidos ao escrever um código com a utilização de macros e como resolver os problemas a partir de exemplos mostrados durante a sessão.

Breakout Session 2857. Base SAS and SAS Enterprise Guide: Automate Your SAS World with Dynamic Code

41. Na apresentação foram discutidas formas de transformar um código estático em um código dinâmico para diminuir os riscos de erro durante a programação, economizar tempo na execução das tarefas ou retrabalho e automatizar os procedimentos. A partir das capacidades de macro da linguagem, da declaração FILENAME, da opção SET INDSNAME e do comando CALL EXECUTE foram apresentados exemplos de código para automatizar rotinas.

Breakout Session 1690-2018. Top 10 Tips for SAS® Enterprise Miner™ Based on 20 Years' Experience

42. A apresentação elencou as 10 melhores práticas relativas ao uso do *SAS Enterprise Miner*, com vistas a aumentar a produtividade na elaboração de modelos de preditivos e ajudar os iniciantes na tecnologia a extrair informação relevante dos dados. A apresentação abordou funcionalidades existentes no SAS 9.4, bem como a recente tendência de integração com Python por meio da plataforma SAS Viya.

Breakout Session 1777-2018. Using SAS Visual Investigator to Enforce Model Tuning Best Practices in a Regulatory Environment

43. O *SAS Visual Investigator* é uma ferramenta que permite, a partir de alertas, analisar as entidades e ações relacionadas ao alerta em busca de informações sobre operações fraudulentas. As regras de alerta devem ser definidas pela instituição e essa apresentação fala sobre como essas regras podem ser melhoradas e ajustadas a partir do feedback do próprio *SAS Visual Investigator*.

44. O procedimento inclui a seleção aleatória de casos que não quebram as regras definidas pelos alertas, mas se aproximam desses limites, e o envio para análise dos investigadores através da ferramenta *SAS Visual Investigator*. A própria ferramenta faz o controle do processo desde a criação do alerta, sua análise e a avaliação final do investigador, mantendo registro das operações realizadas. Posteriormente, o resultado dessa análise pode embasar alterações nos limites das regras que geram alertas, de maneira a diminuir o número de falsos negativos.

Breakout Session 2466-2018. Efficiently Join a SAS® Data Set with External Database Tables

45. A sessão apresentou alguns dos possíveis métodos de realizar junção de bases do SAS com bases externas, avaliando a eficiência e desvantagens de cada um deles. Houve um foco principalmente em bases externas de bancos de dados, como Oracle e SQL Server, que são as bases com as quais normalmente são executadas as junções pela SecexPrevi.

Student Symposium 3602-2018. Bridging the Skill Gap between Post-Secondary Education Outcomes and Employment Opportunities

46. O trabalho dos apresentadores buscou avaliar se as habilidades ensinadas nos cursos de graduação se adequam com as habilidades exigidas pelos empregadores. A informação sobre as habilidades ensinadas nos cursos de graduação foi obtida a partir de mineração de texto nos sites das próprias universidades, que são obrigadas por leis a seguir a ementa divulgada dos cursos. As informações sobre as habilidades requeridas pelos empregadores foi obtida a partir da base de dados de uma empresa de intermediação de mão de obra. A partir desses dados foi construído um relacionamento entre os cursos e vagas de emprego, buscando verificar quais habilidades estariam faltando para a obtenção do emprego pelo graduando.

Breakout Session 1951-2018. Development of an Individual Level Social Determinants of Health (SDoH) Model

47. O trabalho utiliza o CID-10, particularmente as classificações Z55 a Z65 intituladas “Pessoas com riscos potenciais à saúde relacionados com circunstâncias socioeconômicas e psicossociais”. A partir de uma base de atendimento médico, que inclui, além da classificação da condição do paciente, os procedimentos adotados e o custo dos serviços médicos prestados, o trabalho buscou relacionar as diversas condições socioeconômicas com os gastos incorridos nos tratamentos de saúde. Buscou-se avaliar como as condições da população em risco socioeconômico acaba por afetar os gastos com saúde pública do Estado.

Breakout Session 1955-2018. Finding the Treasure: Using Geospatial Data for Better

Results with SAS® Visual Analytics

48. A apresentação mostra as possibilidades de uso de dados georreferenciados no SAS *Visual Analytics*, e as diversas maneiras que esses dados podem ser apresentados: polígonos das áreas, pontos ou esferas no mapa. Também são discutidas as vantagens e desvantagens de cada maneira de apresentar os dados. Por fim, são mostradas algumas ferramentas de obtenção de geolocalização a partir do endereço, tanto por meio de dados da SAS quanto por ferramentas externas, como o API do Google.

Breakout Session 2346-2018. Wow! You Did That Map With SAS®?! Round II

49. Nessa apresentação as capacidades de desenhar mapas com o SAS são apresentadas mais a fundo. Além da possibilidade de criar mapas com áreas coloridas, pontos e esferas, também é apresentada a possibilidade de incluir uma componente temporal nos mapas, possibilitando apresentar uma evolução dos indicadores.

Breakout Session 2141-2018. Location Matters: Evidence from Spatial Econometric Analysis of Opioid Prescribing Rates

50. O trabalho apresenta uma análise geoespacial realizada com beneficiários de programas do governo para compra de remédios e que têm indícios de uso abusivo de substâncias controladas. Foi utilizado o SAS *Visual Analytics* para observar no mapa a distribuição de usuários dos programas com indícios de abusos, assim como os postos de saúde que passaram as prescrições com dosagem elevada e as farmácias que venderam os remédios. A partir dessa visualização espacial foi possível identificar áreas de maior alto risco para o abuso de remédios, auxiliando o direcionamento de políticas públicas para coibir esse problema.

Breakout Session 2684-2018. Macro Method to use Google Maps™ and SAS® to Geocode a Location by Name or Address

51. O método proposto nessa apresentação possibilita que um programador SAS® v9.4 use o Google™ e SAS® para capturar latitude e longitude de qualquer localidade que possa ser identificada por uma pesquisa no Google Maps™. O Google Maps™ é uma aplicação provida gratuitamente pela Google™ muito útil para obter caminhos em mapas além de possibilitar a representação geográfica de endereços (*geocoding*). Em suma, o método proposto possibilita associar endereços existentes em bases de dados estruturadas à uma coordenada geográfica, o que viabiliza uma série de aplicações práticas no TCU.

Breakout Session 2832-2018. Image Classification Using SAS® Enterprise Miner 14.1

52. O trabalho propõe uma abordagem de classificação de imagens usando Python, que pode ser utilizada em ambientes com SAS Viya®. A tarefa de classificação de imagens possui uma série de aplicações que incluem reconhecimento facial, reconhecimento de caligrafia e até análise de imagens de satélite. A aplicação que foi apresentada diz respeito à classificação de caracteres e o modelo foi treinado com uma base 80 mil imagens. Todo o modelo foi desenvolvido usando Python. Porém, com a recente incorporação da linguagem à plataforma Viya, agora pode ser executado no ambiente SAS®.

53. No TCU, existem diversas possibilidades de aplicação de soluções de visão computacional, como, por exemplo, a avaliação de políticas ambientais de combate ao desmatamento ou de políticas de combate à pobreza, ambas por meio de análise de imagens de satélite.

Breakout Session 1868-2018 Robust Tuning for Machine Learning

54. Com o surgimento de algoritmos cada vez mais parametrizáveis e, em especial, as redes neurais, a atividade calibração ou *tuning* de modelos ganhou extrema importância. Modelos de *Machine Learning* podem ter dezenas de hiperparâmetros cuja calibração impacta dramaticamente a performance do processo de modelagem. A apresentação explorou a questão da calibração de modelos aplicada ao Enterprise Miner e mostrou o impacto sobre os resultados do modelo.

Breakout Session 1783-2018. General SAS® Macro to Implement Optimal N:1 Propensity Score Matching within a Maximum Radius

55. *Prompensity Score Matching* (PSM) é um método de inferência causal bastante utilizado para avaliação de efeitos de tratamentos em geral. PSM é aplicado em áreas como medicina, na avaliação de terapias e na econometria, para avaliação de resultados de intervenções, a exemplo da implantação de um programa social. A apresentação mostrou uma abordagem de aplicação do método PSM usando SAS. Métodos que possibilitem avaliar políticas públicas são de grande interesse para o TCU. Com o material apresentado fica claro que o TCU pode usar o SAS para realizar este tipo de análise em seus processos de controle externo.

Breakout Session (Executive). Fraud Detection and Prevention: Key Issues in Health Care

56. Foi mostrado o trabalho realizado pelo *Georgia Department of Public Health* para tentar detectar fraude no uso dos cupons de alimentos distribuídos pelo *Georgia Women, Infants and Children Program*. A partir de diversas tipologias já desenvolvidas para detectar fraudes no programa de cupons, foi usado o *SAS Visual Analytics* para gerar relatórios para cada uma dessas tipologias, apresentando as informações sobre as pessoas e clínicas com problemas em mapas. Essas diversas informações também foram reunidas para permitir uma análise conjunta dos diversos fatores de risco, permitindo descobrir novos padrões de fraude. A partir desses dados eles conseguiram mais que dobrar o número de clínicas investigadas por possíveis fraudes na comercialização dos cupons de alimentos.

Breakout Session (Executive). Winning the Big Data War Pays Big Dividends for Wargaming

57. A apresentação realizada pela empresa de jogos on-line Wargaming apontou para o uso de ferramentas de Análise de dados para maximizar a rentabilidade de seus produtos ao avaliar características da “jogabilidade” de seus jogos. Por meio de mapas de calor, análise de questionários disponibilizados aos usuários, algoritmos de detecção de fraude, modelos preditivos, entre outros, a empresa conseguiu implantar uma cultura de análise de dados nos ambientes de tomada de decisão, permitindo maior planejamento, menor custo de desenvolvimento de novas técnicas e ampliação do alcance de *Analytics*.

Breakout Session (Executive). Could Algorithmic Maturity Be Your Organization’s ‘Secret Sauce’

58. A Syneos Health, multinacional americana que trabalha com pesquisas de estudos clínicos, trouxe a experiência que está tendo na evolução da Análise de dados em sua cultura organizacional. Apontaram a dificuldade em lidar com grandes volumes de dados, provenientes de diferentes fontes em variados graus de credibilidade; as ferramentas e *softwares* que têm utilizado nesse tema; e o caminho e expectativas da Análise de dados na companhia. Destaca-se a clara percepção dos custos e riscos de se implementar um ambiente maduro de *Analytics*, como não haver o efeito “*plug-and-play*” ou a consciência de não haver “super-heróis” entre os especialistas em Análise de dados.

Breakout Session (Executive). Preventing Tragic Outcomes: Using Predictive Analytics in Child Welfare

59. O estado americano da Carolina do Norte está passando por um crescimento no número de casos de crianças em situação de abuso ou negligência, com alta correlação com a crise de derivados do ópio. Com o uso de análise de dados, o condado de New Hanover conseguiu aumentar o acesso e qualidade dos dados de Assistência Social, permitindo, por exemplo, dar a um juiz que decide sobre a guarda de uma criança ou adolescente informações sobre casos de violência reportados à polícia ou se os futuros guardiões teriam saído recentemente do sistema prisional. Apresentaram um sistema de alerta, com notas de risco, a ser utilizado por agentes públicos com poder de decisão nesses casos.

Breakout Session 1683-2018. Data Management Meets Machine Learning

60. A empresa ThatWave Technologies trouxe sua visão sobre a relação entre o aprendizado de máquina (campo de estudos relacionado à Inteligência Artificial, ou *Machine Learning*) e o gerenciamento de dados, como um todo (*Data Management*). Alguns exemplos do uso de Analytics, em casos práticos, incluem: a organização de domínios e dicionários de dados (como trabalhar com tabelas totalmente novas); como recomendar potenciais novas fontes de dados; e a detecção de anomalias.

ENCAMINHAMENTOS POSSÍVEIS, NO ÂMBITO DO TCU, DECORRENTES DESTA AÇÃO

61. Em relação a forma como o software SAS acessa suas bases de dados ou as informações em Sistemas Gerenciadores de Bancos de Dados, como o Labcontas, os conhecimentos adquiridos permitem compreender como ocorrem esses relacionamentos e como será possível otimizar os processos em produção e os processos que serão desenvolvidos com o objetivo de otimizar a performance e os tempos de processamento.

62. Também se pôde observar que muitas análises foram incrementadas através da ferramenta *SAS Visual Analytics*, principalmente com a utilização de mapas e geolocalização das entidades analisadas (pessoas, empresas, ocorrências de eventos, etc.). A ferramenta já foi adquirida pelo TCU e pretende-se avaliar como adaptar as metodologias apresentadas no evento para incrementar as bases já utilizadas pelo Tribunal com informação de geolocalização.

63. Em relação à nova versão do software SAS, nome fantasia “SAS Viya”, os conhecimentos adquiridos permitem identificar as principais diferenças e limitações entre as versões SAS 9.4 e SAS Viya. Conforme observado no evento, as duas versões do software continuam em aprimoramento paralelo, apesar de a empresa estar adaptando novas funcionalidades da versão SAS Viya para a versão SAS 9.4.

64. Destaca-se, contudo, que se identificou o risco de uma maior evolução ocorrer apenas em uma das plataformas (SAS Viya), o que – aliado a um maior uso das ferramentas SAS pelo TCU – pode criar a necessidade de se abandonar a utilização da versão SAS 9.4 para uma migração a uma plataforma integralmente SAS Viya.

65. Adicionalmente, paralelamente à introdução da nova plataforma SAS Viya, o SAS 9.4 está incorporando suporte a linguagens de programação aberta como o Python, o que facilita o desenvolvimento de soluções que envolvam o poder de processamento de grandes volumes de dados do SAS e a utilização de algoritmos de estado da arte lançados em código aberto.

66. Em relação a temática das sessões assistidas e das soluções apresentadas, observou-se que na área de governo foram realizadas algumas apresentações sobre fraude em benefícios ou avaliação de políticas nos temas de Saúde e Educação. Verifica-se a oportunidade para o desenvolvimento de soluções utilizando análise de dados com o SAS dentro dessas áreas e para a obtenção de novos conhecimentos com a participação em futuros fóruns.

67. Além de pretender utilizar as tecnologias aqui listadas nos processos de Fiscalização Contínua de Benefícios no âmbito da SecexPrevidência, como forma de disseminação de conhecimento para outras unidades do TCU, a equipe que participou do evento se coloca à disposição para realização de apresentações tendo como conteúdo a visão geral sobre os conhecimentos adquiridos ou sobre temas específicos de interesse de outras secretarias.

SecexPrevidência, em 23 de abril 2018

Aloísio Dourado Neto

AUFC – Mat. 8814-5

Giuseppe de Abreu Antonaci

AUFC – Mat. 10.695-0

Luiz Henrique Batistuta Gomide

AUFC – Mat. 9981-3

Rodrigo Otávio Coelho Hildebrand

AUFC – Mat. 9466-8

(Assinado eletronicamente)