

Machine Learning Nanodegree

Projeto Final

Michel P. Fernandes

16 de Abril de 2018

Definição

Visão geral do problema

O avanço da sociedade nas áreas de saneamento, medicina, educação, alimentação dentre outras áreas, está ocasionando um crescimento acelerado da população idosa no mundo, aumentando a expectativa de vida das pessoas.

No Brasil, em 1900, segundo dados do IBGE (Instituto Brasileiro de Geografia e Estatística¹), a expectativa de vida ao nascer do brasileiro era de 33,7 anos. Já em 2010 essa expectativa passou para 73,4 anos e a projeção para 2041 é de 80,09 anos. Em função destes dados apontarem para uma população idosa em grande número, Orestes e Jorge (FORLENZA; ALMEIDA, 2006) mostraram que a área dos transtornos neurológicos brasileiros que estão associados ao envelhecimento representa um campo de grande importância para a saúde pública nacional, especialmente sobre os custos envolvidos na assistência das doenças.

Devido a esta alta incidência dos casos de demência, os custos para atender a doença cresce a cada ano. Comparativamente, os custos para o tratamento de demência se equivale, se fosse uma empresa, na maior empresa do mundo, quando comparamos em termos a receita anual, a frente da Wal-Mart (414 bilhões de dólares) e da Exxon Mobil (311 bilhões de dólares). Se a doença fosse um país seria a décima oitava economia mundial ficando entre a Turquia e a Indonésia. Estima-se que os custos para o ano de 2010 sejam de 604 bilhões de dólares (Figura 1 e Figura 2).

¹<http://www.ibge.gov.br/series_estatisticas/exibedados.php?idnivel=BR&idserie=POP321>

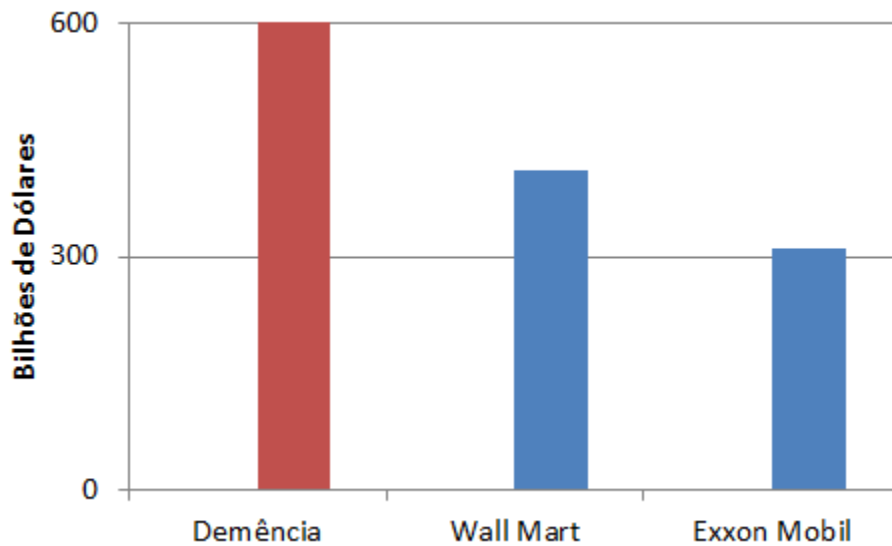


Figura 1: Custo da demência comparado às receitas das grandes empresas. Fonte: Adaptado de *World Alzheimer Report 2010 Alzheimer's Disease International*

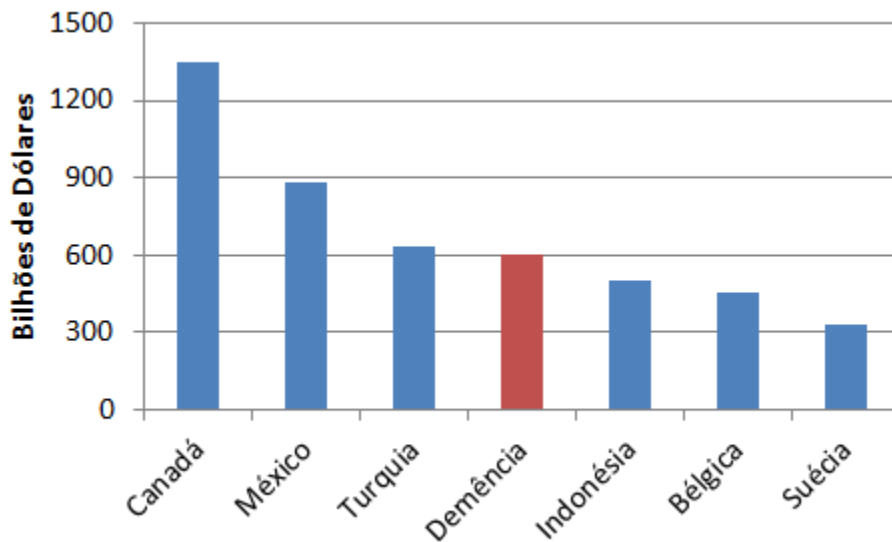


Figura 2: Custo da demência mundial comparado às economias de países. Fonte: Adaptado de *World Alzheimer Report 2010 Alzheimer's Disease International*

Embora os países desenvolvidos tenham a maior quantidade de pacientes com transtornos neurológicos, seu custo em relação ao PIB (Produto Interno Bruto) é menor do que os países em desenvolvimento quando ponderado pela quantidade de pacientes. Por exemplo, o custo da demência para a América do Norte representa 1,06% do PIB (com 4,3 milhões de pacientes) enquanto que na América do Sul tem-se 1,02% (com 0,6 milhões de pacientes), segundo o *World Alzheimer Report 2010 Alzheimer's Disease International*. Esses números mostram o grande impacto na economia para os países em desenvolvimento com o avanço da demência.

Além disso, é verificado que dentre as doenças relacionadas a transtornos neurológicos e a idade avançada, a Doença de Alzheimer (DA) é a que mais prevalece. A medida que a idade avança, a proporção de casos de demência que evolue para DA é cada vez maior, chegando a responder pela quarta parte de todas as demências quando a pessoa acometida tem 85 anos ou mais, conforme ilustrado na tabela .

Idade	Demência	DA
65 anos ou mais	5% a 8%	3% a 4%
75 anos ou mais	12% a 20%	Em torno de 10%
85 anos ou mais	25% a 50%	Em torno de 25%

Tabela 1:
Evolução da DA com a idade.
Fonte: (FORLENZA; ALMEIDA, 2006)

A dificuldade de diagnóstico precoce da DA é o maior desafio da doença, já que se sua identificação for muito tardia o tratamento não poderá contribuir muito para a desaceleração da doença, já que ela ainda não tem cura. O ideal é que ela seja detectada precocemente para que o início dos tratamentos sejam realizados o quanto antes.

Demência

A demência não implica em uma doença específica ou mesmo uma patologia específica, é uma mudança na função cognitiva que é severa o bastante para comprometer as atividades diárias de um indivíduo (BEAL; LANG; LUDOLPH, 2005). Segundo o *Diagnostic and Statistical Manual of Mental Disorders*, quarta edição (DSM-IV), demência pode ser definida como um comprometimento adquirido das funções cognitivas que incluem um declínio de memória além do que é esperado para a idade e pelo menos uma outra função cognitiva, como atenção, habilidades visoespaciais ou linguagem, ou declínio em atividades funcionais como planejamento, organização, sequenciamento ou abstração. O declínio não pode somente afetar habilidades emocionais mas também interferir em atividades profissionais e sociais. Os déficits não devem estar acompanhados pelo comprometimento de excitação (demência) ou verificado por outra condição psiquiátrica, como depressão ou esquizofrenia. Demência pode ser definida também como diagnóstico de etiologia possível, provável ou definido. A demência degenerativa implica a progressão da doença ao longo do tempo.

Embora os critérios do DSM-IV sejam geralmente úteis, um problema com o critério é que o prejuízo da memória é uma característica essencial. Apesar disso ser bastante comum na maioria das demências, em outras pode apresentar comprometimento em domínios que não são de memória. Se a apresentação inicial é uma mudança de personalidade ou comportamento, ao invés de memória, o diagnóstico pode ser Demência Frontotemporal. Em indivíduos com *parkinsonismo*, alucinações e flutuações no comportamento, Demência por Corpos de Lewy pode ser mais provável do que DA. A Demência Vascular (DV) pode envolver alterações abruptas do funcionamento de vasos arteriais com a obstrução dos vasos sanguíneos ou de doença embólica ou pode apresentar insidiosamente se a isquemia subcortical (pequeno derrame cerebral) é responsável pelas alterações na função. A anomia proeminente com outras características do distúrbio de linguagem pode ser um sinal primário de afasia. Uma demência progressiva evoluindo rapidamente de semanas a meses com sintomas psiquiátricos e funções motoras sugere um distúrbio do prião como a Doença de *Creutzfeldt-Jakob* (CJD) (BEAL; LANG; LUDOLPH, 2005).

Resumo Histórico

O entendimento atual da DA evoluiu em três fases (MUNHOZ; FELDMAN, 2002):

- (i) a descrição inicial da DA,
- (ii) a associação entre o grau de patologia e a gravidade da doença e

(iii) os fatores genéticos e avanços bioquímicos.

Dr. Alois Alzheimer publicou um relatório sobre "uma doença comum no córtex cerebral"(STELZMANN; SCHNITZLEIN; MURTAGH, 1995). Ele descreveu o caso de uma mulher com 51 anos, August D. que inicialmente desenvolveu um transtorno delirante porém depois desenvolveu uma "rápida perda de memória". Exames *post-mortem* do seu cérebro revelaram um cérebro atrofiado, sem evidências de degeneração focal. Microscopicamente, ele descreveu a perda de células neuronais e "apenas um emaranhado de fibrilas que indica o lugar onde o neurônio foi previamente localizado"(FORSTL; LEVY, 1991; STELZMANN; SCHNITZLEIN; MURTAGH, 1995; KLUNEMANN et al., 2002).

Na década de 1960, Blessed et al. (BLESSED; TOMLINSON; ROTH, 1968) relatou a associação entre o número de placas senis nas seções do córtex cerebral e declínio em medidas de funcionamento intelectual entre os idosos. Ele também reconheceu que a DA foi a causa mais comum de demência.

Os avanços genéticos nos dias de hoje ajudaram a entender melhor o conhecimento atual da DA. As mutações da proteína precursora β -amilóide no cromossomo 21 e os genes *presenilina* 1 e 2 foram associados como familiar autossômica dominante de DA (SHERRINGTON; ROGAEV; LIANG, 1995; GEORGE-HYSLOP et al., 1996). A herança do alelo ϵ 4 da *apolipoproteína* E (ApoE) é considerada como um fator de risco para a herança de início tardio e/ou patogênese da DA (CORDER; SAUNDERS; STRITTMATTER, 1993). Além disso, podem haver outros contribuintes genéticos a DA como o cromossomo 10 (ERTEKIN-TANER; GRAFF-RADFORD; YOUNKIN, 2000).

Aspectos Clínicos Associados

A evolução clínica da DA pode ser dividida em três categorias:

- (i) insuficiência progressiva da memória,
- (ii) progressiva disfunção cortical (*afasia*², *apraxia*, disfunção visoespacial) e
- (iii) distúrbios neuropsiquiátricos.

O quadro clínico típico da doença é o início insidioso onde em seu início apresenta-se com sinal de benignidade, quando só é manifestado os sintomas quando a doença já evoluiu, e progressão gradual do distúrbio de memória amnésica com dificuldades de aprendizagem e memória. Nos estágios iniciais da doença, o comprometimento da memória geralmente é para informações recém adquiridas com preservação relativa de eventos remotos (BEAL; LANG; LUDOLPH, 2005).

O paciente é geralmente inconsciente de memória ou comprometimento cognitivo (*anosognosia*). O comprometimento da memória sozinho, porém, não é suficiente para fazer o diagnóstico de demência. Segundo os critérios da NINCDS-ARDA (*National Institute of Neurologic, Communicative Disorders and Stroke/Alzheimers Disease and Related Disorders Association*) exige-se que duas ou mais outras esferas de cognição sejam comprometidas, tais como *apraxia*, *afasia* e *agnosia* (MCKHANN et al., 1984). O diagnóstico de demência é realizado na base do cumprimento dos critérios clínicos e como tal, constitui um diagnóstico positivo, em vez de exclusão (SMALL; RABINS; BARRY, 1997). A correlação clínico-patológica é realmente muito alta, na faixa de 80% a 90% (GALASKO; HANSEN; KATZMAN, 1994).

Durante a evolução da doença os pacientes geralmente desenvolvem distúrbio progressivo de linguagem com *anomia*³ e pode progredir para um quadro de fluente *afasia*. Os pacientes podem ser incapazes de realizar mais que uma atividade ao mesmo tempo ou executar tarefas complexas de rastreamento mental bem como ter outros domínios cognitivos comprometidos. Isto pode estar relacionado com a diminuição de concentração, dificuldade com aritmética mental ou tendência a se tornar confuso. Os pacientes se tornam menos capazes de realizar tarefas exigentes da vida diária, tais como a gestão de finanças e condução de veículos, perturbação do raciocínio

²A afasia é a perda da capacidade e das habilidades de linguagem falada e escrita. (Fonte: <<http://www.abcdasaude.com.br/artigo.php?13>> Acessado em 23/09/2010)

³Anomia é a dificuldade em encontrar palavras. (Fonte: <<http://www.serafimcarvalho.net/sm01.asp?idp=4>> Acessado em 18/10/2010)

abstrato, disfunção executiva (percepção, planejamento e julgamento), e perturbações de habilidades visoespaciais (manifestado através de desorientação geográfica ou ambiental e dificuldade de copiar figuras) (BEAL; LANG; LUDOLPH, 2005).

Todos os déficits cognitivos pioram com o tempo. O paciente vai se tornando progressivamente mais dependente de cuidados especiais, principalmente para alimentação e higiene. O *platô* da doença pode ocorrer quando o comprometimento cognitivo não se altera por um período de 1-2 anos, mas a retomada da progressão ocorre (BEAL; LANG; LUDOLPH, 2005).

Sintomas neuropsiquiátricos são comuns em pacientes com DA (CHUNG; L., 2000). Existem quatro principais grupos de sintomas neuropsiquiátricos na DA:

- (i) perturbação do humor;
- (ii) delírios e alucinações;
- (iii) mudanças de personalidade;
- (iv) desordens de comportamento.

A mudança de personalidade pode ser a primeira alteração clínica observada na DA e pode ser vista em pelo menos 75% dos pacientes. A maioria dos casos se configura pela apatia que aparece no início da doença com sintomas de desinteresse em geral. Esses sintomas podem estar associadas com a depressão (BEAL; LANG; LUDOLPH, 2005).

Conforme a doença progride, agitação e agressividade podem aparecer. Distúrbio comportamental pode ser observado em 30% a 85% dos pacientes e positivamente correlacionada com demência severa. Dentre todos os sintomas neuropsiquiátricos, este pode ser o mais difícil de gerenciar. Os pacientes podem ser tanto fisicamente quanto verbalmente violentos além de poderem ter incontinência urinária (BEAL; LANG; LUDOLPH, 2005).

Durante a progressão da demência se desenvolvem delírios, alucinações e outros comportamentos psicóticos. São verificados comumente delírios de roubo da mesma forma como alucinações auditivas e visuais. Não reconhecimento de pessoas também podem ocorrer (por exemplo, alegando que o cônjuge é um impostor) e também poderá se perder em ambientes e lugares que antes eram conhecidos (BEAL; LANG; LUDOLPH, 2005).

Os transtornos de humor podem ocorrer a qualquer momento até mesmo sem a doença, especialmente com a depressão. As taxas de depressão podem variar de 25% a 30% com DA provável e é possível encontrar casos onde a taxa pode chegar a 50% (ZUBENKO; ZUBENKO; MCPHERSON, 2003). Tekin e Cummings especulam que a depressão pode estar relacionada com degeneração estrutural e mudanças bioquímicas no cérebro com perda neuronal em *adrenérgicos*, sistemas *colinérgicos* e *serotoninérgicos* (TEKIN; CUMMINGS, 2001).

Alterações motoras estão ausentes na DA até os últimos anos da doença. A presença de anormalidades focais, alterações da maneira de andar, convulsões precoces do curso clínico da DA torna-se improvável (BEAL; LANG; LUDOLPH, 2005).

Após o início da DA, os pacientes sobrevivem de 2 a 20 anos, mas na maioria dos casos a sobrevivência da doença é de 7 a 10 anos. A taxa média anual de declínio reportado no MMSE (*Mini-Mental State Examination*) é de 4 a 5 pontos. Os pacientes normalmente morrem de complicações de saúde como bronquite ou pneumonia (BEAL; LANG; LUDOLPH, 2005).

0.1 Diagnóstico Clínico

O diagnóstico da DA pode ser feito através de critérios do NINCDS-ADRDA (MCKHANN et al., 1984) ou do DSM-IV. A avaliação e o diagnóstico da DA exigem identificação das características principais e clínicas e exclusão de outras causas mais comuns de demência. As diretrizes do relatório do *American Academy of Neurology*, representadas na tabela 2 informam que os critérios do DSM-IV e do NINCDS-ADRDA para o diagnóstico da DA têm confiabilidade e validade e devem ser usados (KNOPMAN; DEKOSKY; CUMMINGS, 2001).

O ponto mais importante para se verificar o diagnóstico de demência é a investigação não somente do histórico do paciente mas também de um acompanhante confiável. Informações para a obtenção de histórico

O diagnóstico clínico de provável doença de Alzheimer
Demência estabelecida por exame clínico e mental e confirmado por testes neuropsicológicos
Déficits em pelo menos dois domínios cognitivos
Declínio cognitivo progressivo, incluindo a memória
Nível normal de consciência
Início entre as idades de 40 e 90 (o mais comum após os 65 anos)
Não há explicação médica ou neurológica possível
O diagnóstico provável da doença de Alzheimer é apoiado por
Afasia progressiva, apraxia e agnosia
Atividades da vida diária prejudicadas
Histórico familiar de doença semelhante
A atrofia do cérebro na tomografia computadorizada ou ressonância magnética, especialmente se progressiva
CSF e EEG (Eletroencefalograma) normais ou inespecíficos
Outras características clínicas consistentes com provável doença de Alzheimer
<i>Platô em curso</i>
Sintomas: depressão, insônia, incontinência, ilusões, alucinações, reação verbal, emocional ou física violenta, distúrbios sexuais, perda de peso, durante os estágios avançados observa o aumento do tônus muscular, <i>mioclonia</i> e caminhar anormal
Convulsões na doença avançada
Tomografia computadorizada normal para a idade
Características que fazem a doença de Alzheimer incerta ou improvável
Início agudo
Sinais focais sensório-motor
Convulsões ou distúrbios ao andar precoce
O diagnóstico clínico da doença de Alzheimer possível
Demência atipicamente precoce ou em curso na ausência de outra explicação médica ou neuropsiquiátrica
Demência com outra doença, não percebido como ser a causa da demência
Para efeitos de pesquisa, um déficit focal progressivo cognitivo
Doença de Alzheimer Definitivo
Cumprir os critérios clínicos para a doença de Alzheimer provável
Tecido de confirmação (autópsia ou biópsia do cérebro)
Pesquisa de classificação da doença de Alzheimer deve especificar
Familiar?
O início precoce (antes dos 65 anos)?
Síndrome de Down (<i>trisomia 21</i>)?
Coexistente em outras doenças neurodegenerativas?

Tabela 2:

Critério do *National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's and Related Diseases Association* (NINCDS-ARDA).

(Adaptado de (MCKHANN et al., 1984))

incluem uma mudança e um declínio no nível de habilidade do paciente para executar as decisões, planejar (refeições por exemplo), gerir finanças (contas, talão de cheques, pagamento de impostos), e de dirigir (por exemplo, se perder durante a condução de veículos, acidentes de trânsito). Também é pertinente uma história do paciente se repetir. Embora existam vários instrumentos que podem ser utilizados para obter essa informação, pode-se indagar sobre a capacidade do doente e dificuldades para realizar atividades típicas do cotidiano (BEAL; LANG; LUDOLPH, 2005).

Além do histórico do paciente, uma avaliação cognitiva como um instrumento para avaliar a função mental pode ser bastante útil, especialmente para realizar o acompanhamento do paciente. Instrumentos para esse fim incluem o *Mini-Mental State Examination* (MMSE) (FOLSTEIN; FOLSTEIN; MCHUGH, 1975), o *Modified Mini-Mental State* (3MS) (TENG; CHUI, 1987), o *Blessed Orientation Memory Concentration Test* (BOMC) (KATZMAN et al., 1983) e o *Short Test of Mental Status* (STMS) (KOKMEN et al., 1991). A entrevista estruturada fornecida pelo *Clinical Dementia Rating* (Taxa de Demência Clínica) (CDR) (MORRIS, 1993) é útil

para caracterizar ainda mais a capacidade do paciente demente em vários domínios cognitivos e de desempenho funcional. Embora o CDR possa ser útil, não é superior aos outros.

Indivíduos que estão sendo avaliados para demência devem também ter um exame neurológico geral. Normalmente, no início da DA, esse exame é bastante normal, com exceção da avaliação do estado mental. No entanto, no decorrer do exame, outras características podem sugerir outros fatores para que a demência possa ser explicada. Por exemplo, descobertas de *parkinsonismo* pode sugerir um componente de corpos de *Lewy* ou a presença de reflexos assimétricos ou corte visual de campo ou outros sinais podem sugerir um componente vascular. Da mesma forma, se há outras características neurológicas como uma neuropatia periférica, estas podem sugerir problemas tóxicos ou metabólicos. A avaliação da função sensorial é importante uma vez que podem afetar o estado mental e o exame neurológico. Finalmente, o exame neurológico deve ser complementado por um exame de um médico generalista procurando outras contribuições sistêmicas para o prejuízo cognitivo (BEAL; LANG; LUDOLPH, 2005).

0.1.1 Exame do Fluido Espinhal

Há pouca evidência para recomendar o uso rotineiro da punção lombar na avaliação de rotina de pacientes idosos da demência (Becker et al., 1985). No entanto, na ausência de contra-indicações, pode haver certas circunstâncias clínicas em que o exame do líquido espinhal é recomendado. O médico pode suspeitar de um processo alternativo contribuindo nas demências caracterizado por uma alteração do estado mental subaguda, *encefalopatia* crônica, apresentação clínica incomum, início precoce da doença (menor que 65 anos), presença de febre e/ou rigidez de nuca, câncer sistêmico, pacientes imunocomprometidos ou doenças do colágeno. Se o exame do líquido espinhal é realizado, além de exames de rotina, tais como proteínas totais, glicose, hemograma, citologia, sorologia para sífilis, deverá também considerar as culturas para bactérias, fungos e microbactérias além de índices imunológicos, como índice de IgG (*Imunoglobulina G*), a taxa de síntese de IgG e Bandas Oligoclonais (BO). A positividade de qualquer um dos índices imunológicos indica inflamação e não DA. Pressão de abertura também deve ser registrada principalmente em pacientes com suspeita de hidrocefalia comunicante.

0.1.2 Neuroimagem

Neuroimagem estrutural com tomografia computadorizada sem contraste ou ressonância magnética é recomendado na rotina de avaliação de pacientes com demência e pode ser útil na exclusão de causas reversíveis e tratáveis de demência (KNOPMAN; DEKOSKY; CUMMINGS, 2001). Apreciação do hematoma subdural, neoplasias, infarto encefalomalacia traumática anterior, atrofia focal relevante e realce meningeal daria um diagnóstico diferente de DA. Atrofia das estruturas do lobo bilateral temporal medial, tais como o hipocampo, foi encontrada em pacientes com DA (Figura 3) (JACK; PETERSEN; XU, 1997). No entanto, esta atrofia pode ser inespecífica e pode ser vista em outras condições, mas é certamente consistente com DA e é um marcador bastante sensível da fase patológica e conseqüente estado cognitivo (FOX et al., 1999). Os dados recentes sobre medições longitudinais volumétricas do hipocampo também indicaram que a taxa de progressão da atrofia das pessoas com DA é superior ao de controles normais (JACK; PETERSEN; XU, 1999) e pode preceder os sintomas clínicos de declínio cognitivo (FOX; WARRINGTON; FREEBOROUGH, 1996).

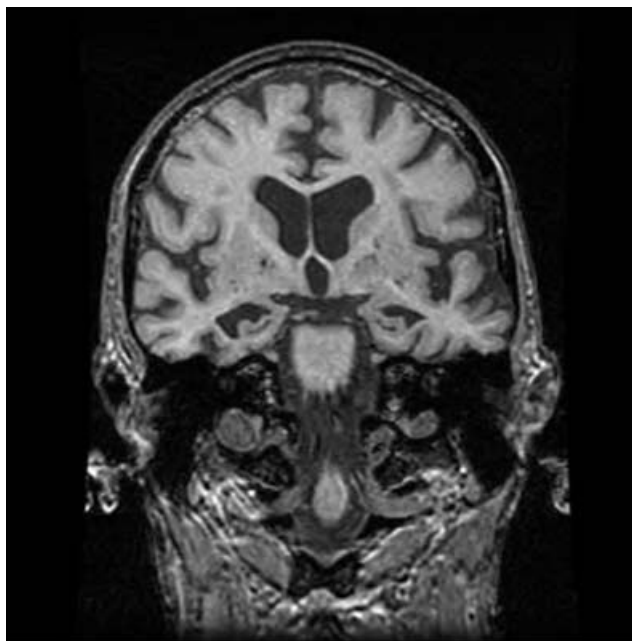


Figura 3: Imagem de Ressonância Magnética Coronal ponderado em T1 mostrando atrofia do hipocampo em um paciente clinicamente com DA provável. Fonte: (BEAL; LANG; LUDOLPH, 2005, p 420)

Imagens funcionais como tomografia por emissão de fóton único (SPECT) ou tomografia por emissão de pósitrons (PET) geralmente revelam hipoperfusão temporoparietal bilateral e do cíngulo posterior e hipometabolismo em DA. Diversos estudos de SPECT recomendam a importância da imagem funcional em aumentar a perspicácia do exame clínico no diagnóstico da DA e pode também se correlacionar com o estado patológico de *Braak*⁴ (VANGOOL et al., 1995; CLAUSS; HARKAMP; BRETELER, 1994). Contudo, com base em estudos Classe II, a sensibilidade do SPECT tem sido menor do que o diagnóstico clínico (KNOPMAN; DEKOSKY; CUMMINGS, 2001).

Exames de imagem funcional podem ser especialmente úteis no diagnóstico diferencial das demências. Em particular, a capacidade de PET e SPECT para diferenciar demência frontotemporal da DA pode ser útil com os temas demência frontotemporal mostrando hipoperfusão predominante e hipometabolismo nos lobos frontais com relativa reserva de estruturas posteriores (BEAL; LANG; LUDOLPH, 2005).

Embora os testes de neuroimagem funcional mostrem um grande potencial, no momento, a *American Academy of Neurology* indicou que nem SPECT ou PET foram recomendadas para uso rotineiro no diagnóstico inicial ou diferencial das demências (KNOPMAN; DEKOSKY; CUMMINGS, 2001).

Descrição do problema

Hoje em dia a análise de imagens médicas é uma atividade que consome um tempo muito grande e seu resultado efetivo nem sempre resultará no padrão completo que se deseja extrair, por exemplo, em uma imagem médica de um paciente com DA se deseja captar o padrão de atrofia de massa cinzenta do cérebro (KLOPPEL et al., 2008).

Alguns métodos de classificação de imagens médicas existem, porém em sua grande maioria são estudos univariados, utilizando somente a informação da imagem médica, descartando outras informações adicionais que poderiam trabalhar de maneira cooperativa para enriquecer o modelo de análise.

⁴Estágios nas alterações patológicas neurofibrilares do cérebro na doença de Alzheimer

Portanto, combinar diferentes atributos fisiológicos, especialmente os que trazem dados a cerca da anatomia cerebral a qual podemos entender como causa, com outros atributos mais objetivos de impacto direto, ou efeito, de como a demência é manifestada no aspecto cognitivo.

Deste modo, foi desenvolvido um modelo de classificação binária capaz de separar adequadamente sujeitos com DA e sujeitos de controle em diferentes momentos de coleta de exames para não apenas validar o modelo de classificação como também comprovar o efeito degenerativo da DA ao longo do tempo.

Métrica

Este estudo tem como objetivo classificar adequadamente sujeitos com DA para futuros acompanhamentos clínicos apropriados. Um dos achados mais relevantes é a capacidade de predição com antecedência, pois desta forma é possível acompanhar melhor a doença e planejar ações com maior antecedência.

Diante disto, o indicador que melhor descreve este objetivo é a sensibilidade do modelo, definido pela fórmula 1 e obtido pela Matriz de Confusão representado na tabela 3.

$$Sensibilidade = \frac{VP}{VP + FN} \quad (1)$$

Resultado	Doença Presente	Doença Ausente
Positivo	VP (Verdadeiro Positivo)	FP (Falso Positivo)
Negativo	FN (Falso Negativo)	VN (Verdadeiro Negativo)

Tabela 3:

Quadro explicativo para o conceito de avaliação dos resultados pelo sistema classificador.

Análise

Exploração dos dados

Os dados deste estudo foram fornecidos pela iniciativa Open Access Series of Imaging Studies (OASIS⁵, a qual contém informações parametrizadas a cerca de exames de ressonância magnética e avaliações de testes de cognição do Mini Mental State Examination (MMSE) e o Clinical Dementia Rating (CDR).

Os dados utilizados são particularmente úteis pois é possível relacionar e combinar os dados de fatores sócio-econômicos que podem ser associados a evolução da DA com os achados neurológicos, por meio do entendimento de certos parâmetros dos exames de ressonância magnética.

Os dois grupos de dados são divididos em :

1. Dados transversais de ressonância magnética em jovens, meia-idade, sem demência e adulto idosos com demência Este conjunto consiste em uma coleta transversal de 416 indivíduos com idade entre 18 e 96 anos. Para cada indivíduo, 3 ou 4 exames de RM ponderada em T1 em sessões únicas estão incluídas. Os sujeitos são todos destros e incluem homens e mulheres. Há 100 indivíduos incluídos com mais de 60 anos foram diagnosticados clinicamente com DA muito leve a moderada. Além disso, um conjunto de dados de confiabilidade está incluído, contendo 20 indivíduos não-contemplados, que foram visualizados em uma visita subsequente dentro de 90 dias da sessão inicial.
2. Dados longitudinais de RM em adultos idosos sem demência e com demência. Este conjunto consiste em uma coleção longitudinal de 150 indivíduos com idades entre 60 e 96 anos. Cada sujeito foi examinado em duas ou mais visitas, separadas por pelo menos um ano para um total de 373 sessões de imagens. Para

⁵Disponível em <<http://www.oasis-brains.org/>>

cada sujeito, 3 ou 4 exames individuais de RM ponderados em T1 obtidos em sessões únicas são incluídos. Os sujeitos são todos destros e incluem homens e mulheres. Há 72 sujeitos foram caracterizados como não tratados ao longo do estudo. Outros 64 indivíduos incluídos foram caracterizados como dementes no momento de suas visitas iniciais e assim permaneceram para exames subsequentes, incluindo 51 indivíduos com DA leve a moderada. Outros 14 indivíduos foram caracterizados como não dementes no momento de sua visita inicial e foram posteriormente caracterizados como dementes em uma visita posterior.

As informações obtidas pelo conjunto de dados permite entender e parametrizar os resultados dos exames de RM nos seguintes atributos: volume intracraniano, volume cerebral normalizado, fator de escala Atlas.

Do ponto de vista de metadados social e estado mental, temos no conjunto de dados os atributos: tempo de educação, estado sócio-econômico, avaliação MMSE Mental e avaliação CDR.

Ao total, para cada sujeito, foi utilizado 7 atributos para classificar o conjunto de dados, sendo 3 fisiológicos e 4 objetivos, sendo 2 de condições educacionais e sócio-econômicas e os demais resultados dos testes de cognição.

De acordo com o tamanho do conjunto de dados e dos benchmarks que desejamos comparar, optamos por utilizar os algoritmos de classificação Random Forest e Support Vector Machines (SVM).

O SVM tem aplicação mais utilizada no ramo de identificação de padrões, sobretudo em imagens. Por exemplo, na área de imagens médicas, possui diversos casos de uso onde é possível identificar padrões que podem indicar evolução para diferentes tipos de enfermidades, seja em imagens de duas dimensões como por exemplo um simples raio-x (ÇAMLICA; TIZHOOSH; KHALVATI, 2015) ou até mesmo imagens volumétricas, de três dimensões de uma ressonância magnética (FERNANDES et al., 2012).

A base do SVM é a sua característica principal em maximizar a separação entre duas classes, o que chamamos de uma classificação binária. Esta característica permite escolher, dentre os elementos dos conjuntos de análise no processo de separação, somente os elementos que maximizam a separação entre as classes, que são chamados de vetores de suporte. Com isto, o SVM é robusto e eficiente ao lidar com problemas onde existe uma grande dimensionalidade, seja em atributos ou em exemplos de treinamento, da ordem de milhões. Outro ponto a ser destacado é o uso das funções kernel para separar classes que não sejam linearmente separáveis num mesmo plano. A técnica consiste em criar um novo hiperplano para que seja possível separar as duas classes com uma função linear.

Apesar do SVM ser robusto ao criar um modelo com grande dimensionalidade pode levar mais tempo que outros classificadores, como por exemplo as Árvores de Decisão na etapa de treinamento. A escolha da função kernel apropriada pode levar tempo pois requer maior estudo dos dados envolvidos. Adicionalmente é necessário realizar diversas tarefas de regularização do modelo, portanto sua configuração não é trivial como Naive Bayes ou ainda como as Árvores de Decisão.

O algoritmo de classificação Random Forest é uma agregação de diversas Decision Trees com o objetivo de torná-las mais eficientes e as otimizando.

As Decision Trees, representam um dos mais simples algoritmos, pois é composto de uma série de perguntas e respostas dispostas em forma de árvore binária, ou seja, com apenas duas ramificações.

Sua aplicação está relacionada nas áreas de Discovery Knowledge especialmente nos casos de uso de suporte técnico (SUNDHARI, 2011), sistemas de recomendação (ÇAMLICA; TIZHOOSH; KHALVATI, 2015), sistemas especialistas de seguros (FREMPONG NIMO NICHOLAS, 2017), diagnósticos médicos (TANNER et al., 2008), dentre outros.

A grande vantagem das Decision Trees é sua simplicidade de construção, pois não requer muitos ajustes ou regularizações. Sua forma gráfica de representar o modelo também contribui para se tornar um dos classificadores de mais fácil entendimento.

Independente da separação ser linear ou não linear, o esforço envolvido para construção do modelo é similar, portanto não há maior esforço computacional quando a separação for do tipo não linear.

Uma das limitações mais importantes das Decision Trees é relacionado ao overfitting devido a natureza do classificador, onde a fronteira de decisão tende a ser paralelo aos eixos existentes do plano de amostras. Sem um ajuste de pruning, por exemplo, podem haver distorções na fronteira de decisão que pode prejudicar distribuições suaves. Este efeito também pode ser minimizado com o uso dos ensemble models, como em Random Forest.

Visualização exploratória

Para termos um melhor entendimento da influência e importância de cada atributo disponíveis, optamos por utilizar gráficos quantitativos cuja finalidade é agrupar os atributos em cada sujeito do conjunto de dados.

Nestes gráficos foi avaliado os atributos "M/F", "Age", "EDUC", "SES", "MMSE", "CDR", "eTIV", "nWBV", "ASF". Foram removidos os atributos "MR Delay" pois não indica relação com o estado de demência ou não, é apenas um indicativo na tomada de exame de RM e o atributo "Hand" pois não há variação

O aspecto de sexo associado a demência, neste conjunto de dados, é predominantemente masculino indicado pela figura . De forma oposta, o sexo feminino tem predominância a não demência.

Neste aspecto é importante mencionar que há pesquisas que indicam o oposto, como o estudo de (LIN et al., 2015) sobre o efeitos da demência inicial causado pelo transtorno cognitivo leve. Logo, podemos concluir que o predomínio de homens com demência pode ser um mero efeito de conjunto de dados e não determina uma condição de causa da demência

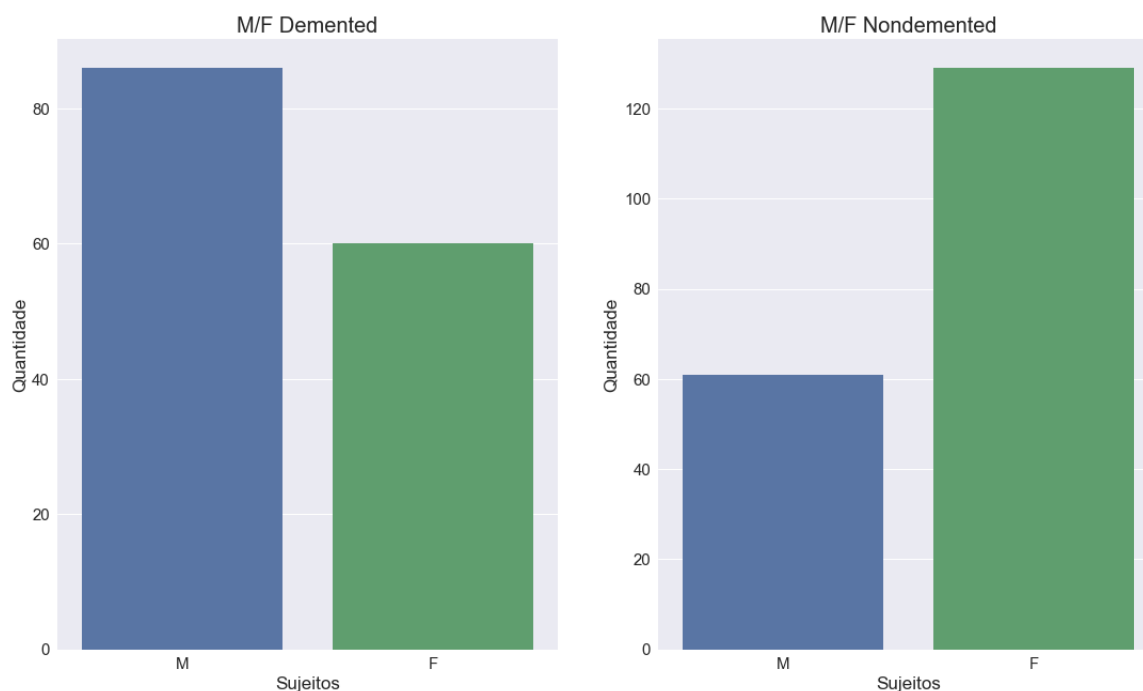


Figura 4: Sexo

A idade (figura não é um fator determinante para diferenciar as duas classes de demência e não demência. Seu impacto é menor, mas significativo pelos aspectos já conhecidos da doença, pois a demência é efeito resultante do envelhecimento da população, sendo considerada uma das condições que é resultado do desenvolvimento humano, conforme discutido nos capítulos introdutórios.

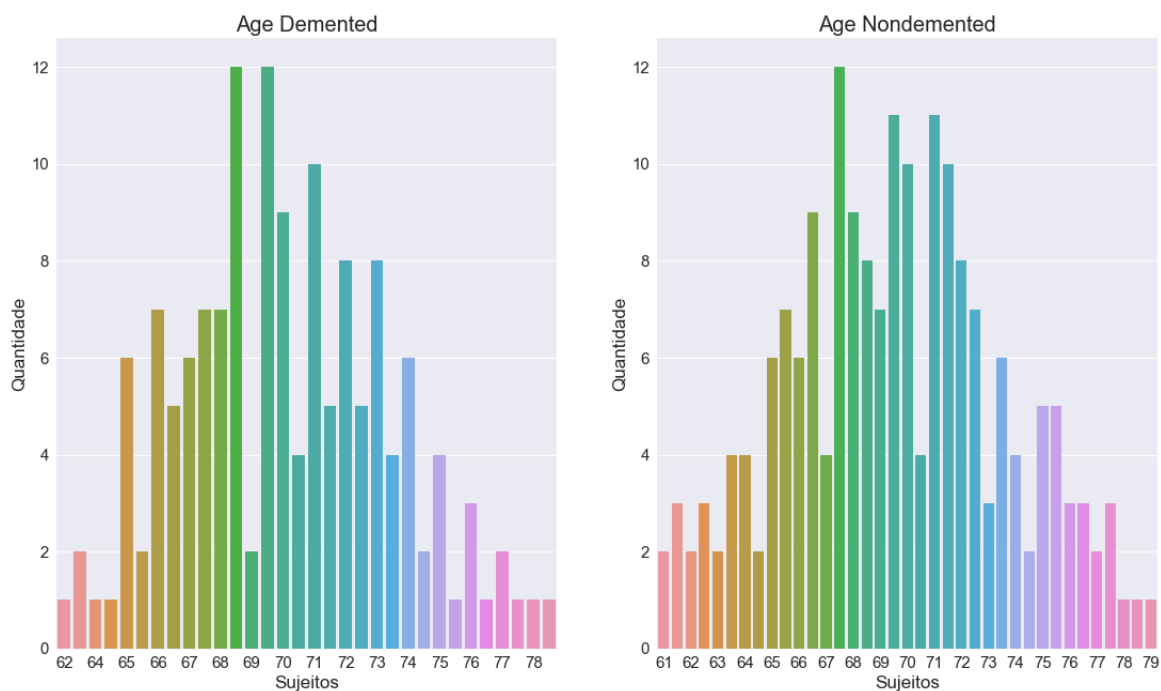


Figura 5: Idade

O nível de educação é caracterizado como uma dos efeitos protetores da demência. Quanto maior o nível educacional, menor é a taxa de conversão.

No gráfico é nítido verificar que no grupo de demência o nível educacional é menor quando comparado com o grupo de não demência.

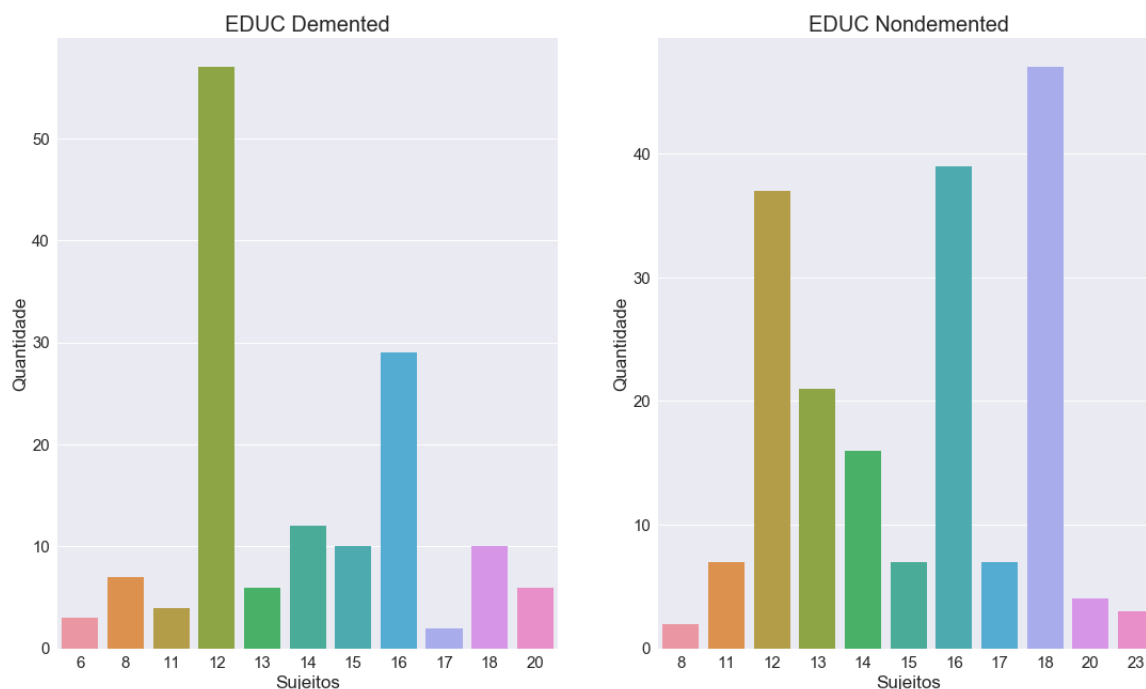


Figura 6: Educação

O gráfico a seguir mostra a divisão dos grupos baseado em índice socio-econômico. Tal índice foi desenvolvido pelo Hollingshead Index of Social Position. As categorias variam de 1 como maior status sócio-econômico e 5 o menor índice.

Para o grupo de demência há mais sujeitos classificados em status de menor índice sócio-econômico, principalmente nos índices 3 e 4. O grupo de não demência apresenta maior número de índices sócio-econômico, especialmente o índice 2.

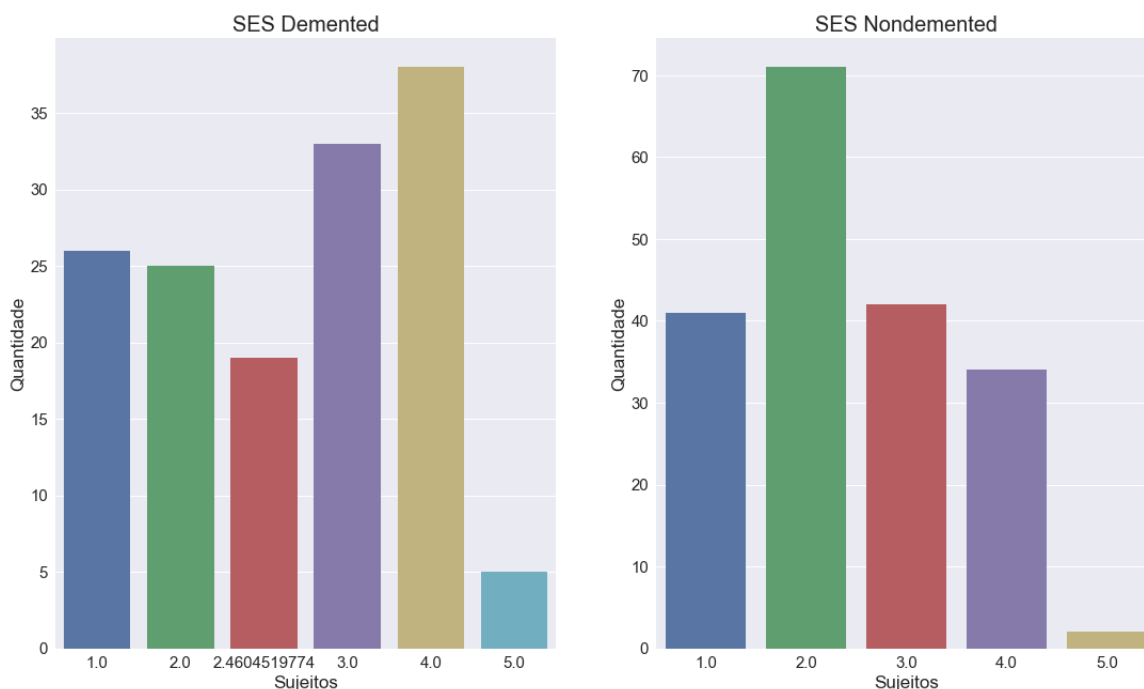


Figura 7: Índice sócio-econômico (SES)

O mini exame do estado mental (ou Mini Mental State Examination, MMSE), exame breve do estado mental é um breve questionário de 30 pontos usado para rastrear perdas cognitivas. É comumente utilizado em medicina para rastrear demência. É também utilizado para estimar a severidade da perda cognitiva em um momento específico e seguir o curso de mudanças cognitivas em um indivíduo através do tempo, portanto fazendo dele um meio efetivo de documentar a resposta do indivíduo ao tratamento.

Qualquer pontuação igual ou superior a 27 (de um total de 30) é efetivamente normal (intacto). Abaixo disso, a pontuação pode indicar perda cognitiva grave (<9 pontos), moderada (10 a 20 pontos) ou leve (21 a 24 pontos) (D., 1991). A pontuação bruta pode precisar ser corrigida de acordo com a escolaridade e idade. Pontuações baixas ou muito baixas são fortemente correlacionadas com demência, embora outros distúrbios mentais podem também levar a resultados anormais no teste. A presença de problemas puramente físicos pode também interferir com a interpretação se não levados em consideração apropriadamente; por exemplo, um paciente pode não ser capaz de ouvir ou ler instruções adequadamente ou pode possuir um déficit motor que afete a habilidade de escrever ou desenhar.

Nos gráficos da é claro a distribuição da pontuação do teste. Para o grupo de demência não há pontuação superior a 18, que indica demência até o limite de moderado. Para o grupo de não demência, há predominância da pontuação 29, o que demonstra ausência de demência completamente.

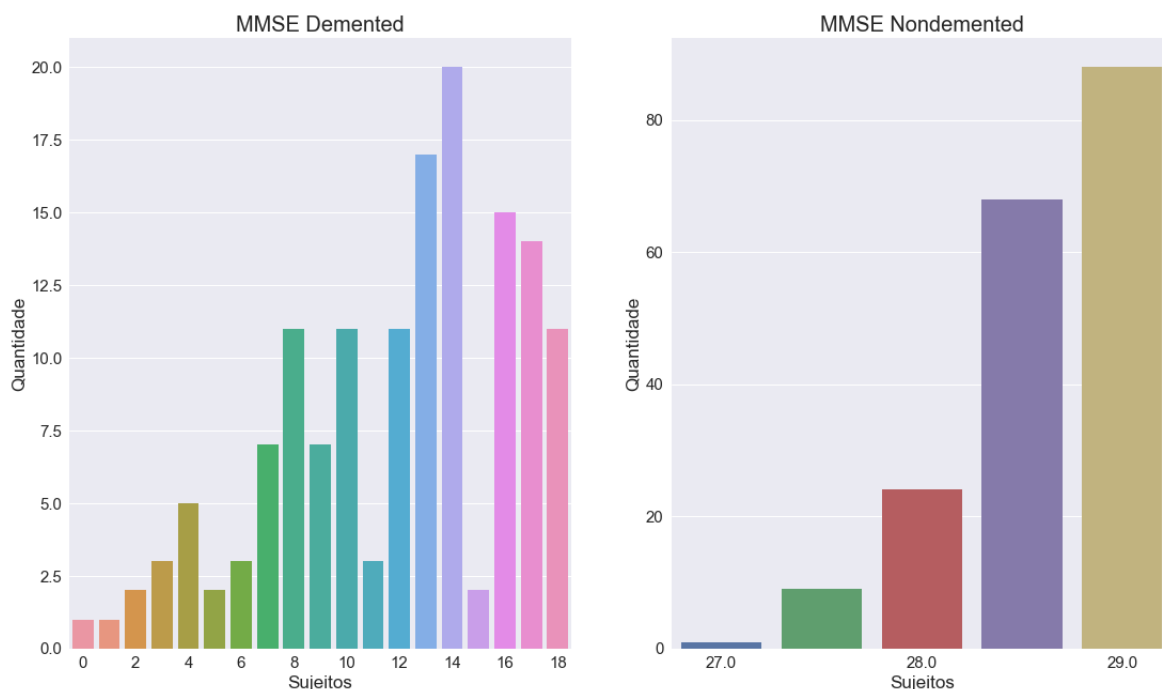


Figura 8: Pontuação do teste MMSE

O teste de CDR é uma escala de 5 pontos usada para caracterizar seis domínios de desempenho cognitivo e funcional aplicáveis à doença de Alzheimer e demências relacionadas como memória, orientação, julgamento e resolução de problemas, assuntos comunitários, casa e passatempos e cuidados pessoais. A informação necessária para fazer cada classificação é obtida através de uma entrevista semi-estruturada do paciente e um informante confiável ou fonte de garantia (por exemplo, membro da família) referido como o Protocolo de Avaliação CDR.

A Tabela de pontuação fornece âncoras descritivas que orientam o médico a fazer avaliações apropriadas com base nos dados da entrevista e no julgamento clínico. Além das classificações para cada domínio, uma pontuação geral do CDR pode ser calculada através do uso de um algoritmo de pontuação CDR. Essa pontuação é útil para caracterizar e rastrear o nível de comprometimento ou demência de um paciente.

- (i) 0: saudável
- (ii) 0,5: demência questionável
- (iii) 1: demência leve
- (iv) 2: demência moderada
- (v) 3: demência grave

A pontuação mostrada nos gráficos a seguir separam bem os grupos de análise. O grupo de demência possui sujeitos a partir da pontuação 0,5, que indica demência questionável indo até demência moderada, segundo a classificação do estudo. Já o grupo de não demência, a totalidade de sujeitos se encontra na pontuação 0, ausência de demência.

Este parâmetro por si só não explica a demência, principalmente por ter um alto grau subjetivo, mesmo apoiando por um teste estruturado. No entanto, quando combinado com outros testes e exames ajudam a construir um diagnóstico mais completo. Por esta razão é que não temos pontuações nos extremos, mesmo para o grupo de sujeitos classificados com demência.

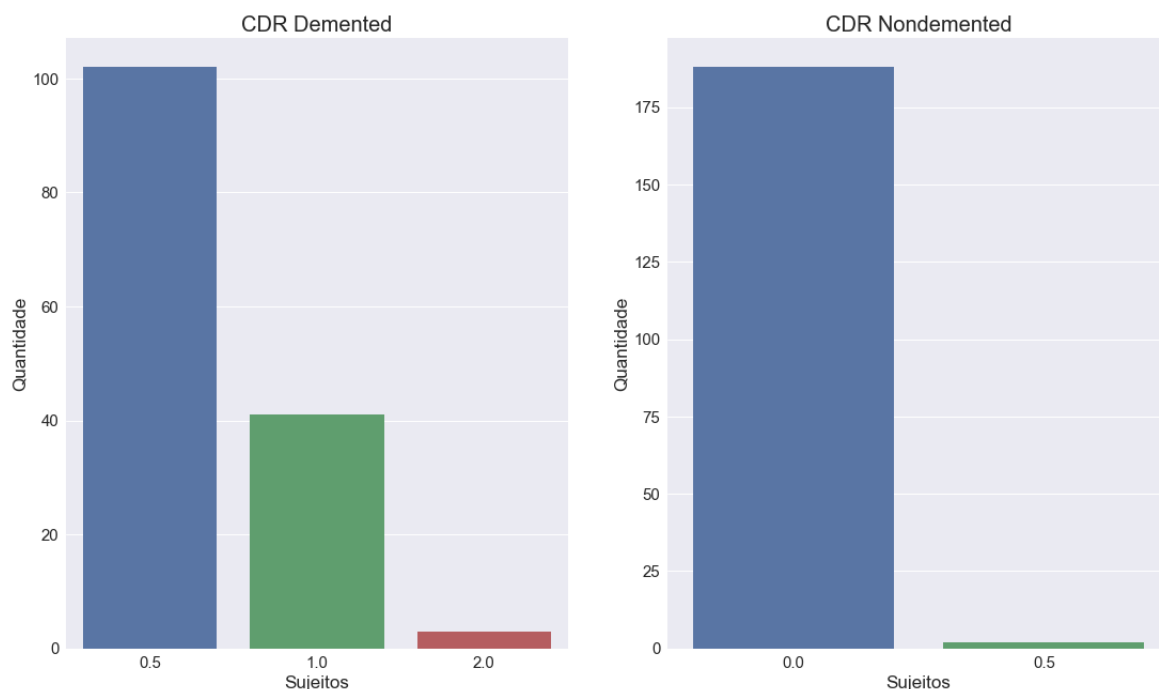


Figura 9: Pontuação do teste CDR

A medida do volume intracraniano total (eTIV), refere-se ao volume estimado da cavidade craniana, conforme descrito pela matéria supratentorial da dura-máter ou pelo contorno cerebral, quando a dura não é claramente detectável. O eTIV é frequentemente usado em estudos envolvidos com a análise da estrutura cerebral sob diferentes modalidades de imagem, como RM, DTI, SPECT, ultrassom e tomografia computadorizada (TC). A consistência do eTIV durante o envelhecimento faz dele uma ferramenta confiável para a correção da variação do tamanho da cabeça em indivíduos que se baseiam em características morfológicas do cérebro. Juntamente com a idade e o sexo, são relatados como covariáveis para ajustar as análises de regressão na investigação de distúrbios cerebrais neurodegenerativos progressivos, como a DA, o envelhecimento e o comprometimento cognitivo. O eTIV também tem sido utilizado como um recurso morfométrico baseado em voxel independente para avaliar mudanças relacionadas à idade na estrutura de braquias pré-mórbidas, determinar padrões de atrofia característicos em indivíduos com TCL e DA, delinear anormalidades estruturais no matéria branca esquizofrenia, epilepsia e eficácia cognitiva.

Os gráficos apresentados são bem específicos e não concentrados, o que não permite uma visualização clara que possa separar os grupos adequadamente.

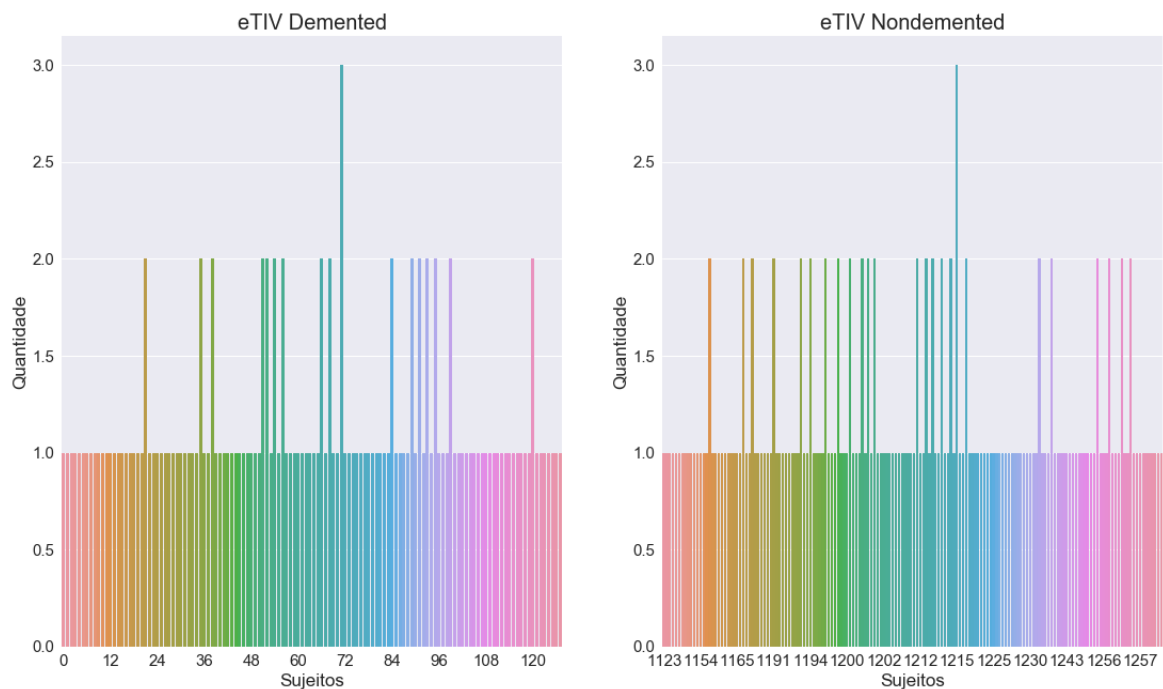


Figura 10: Volume intracraniano estimado

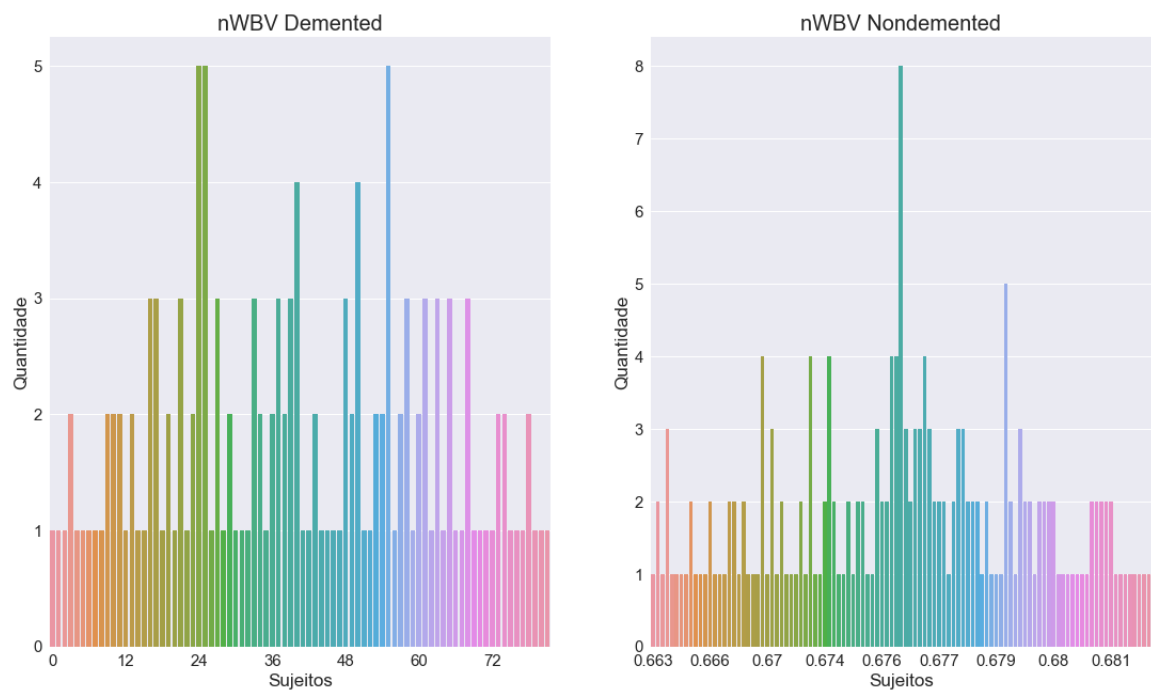


Figura 11: Volume cerebral normalizado

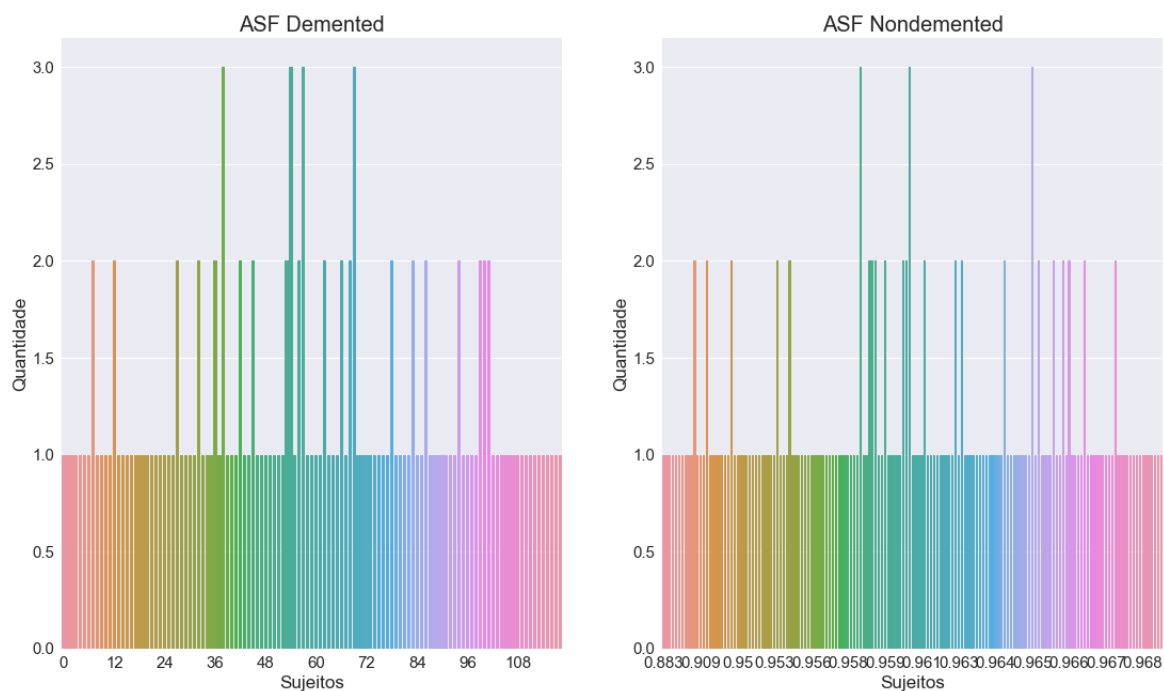


Figura 12: Fator de escala Atlas

Finalmente, foi avaliado a correlação dos atributos do conjunto de dados. No gráfico podemos verificar que dentre os atributos selecionados para construir o classificador, isto é, "M/F", "Age", "EDUC", "SES", "MMSE", "CDR", "eTIV", "nWBV", "ASF" não há correlação direta entre elas, portanto não será necessário remover nenhum atributo destes apresentados.

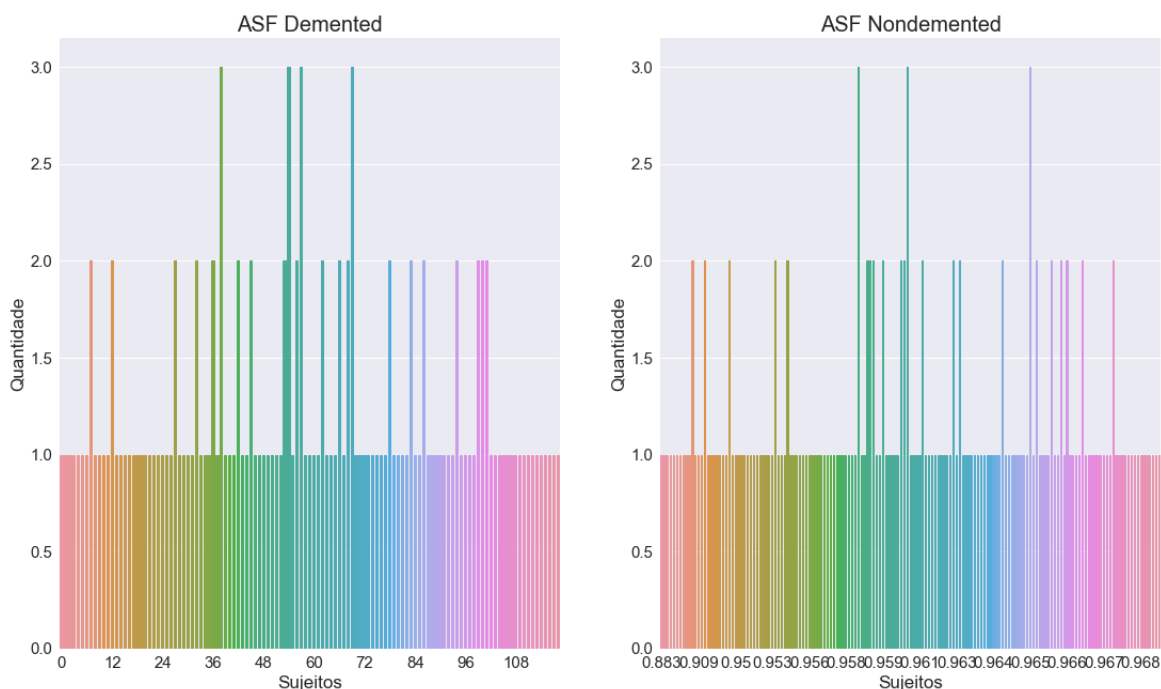


Figura 13: Correlação de atributos

Algoritmos e técnicas

Benchmark

Existem diversos estudos (ZHANG et al., 2015; MORADI et al., 2015) que optaram pelo caminho de análise e processamento diretamente nas imagens médicas ao invés de utilizar os dados obtidos diretamente pelos exames. Tais estudos são mais abrangentes e permitem extrair um número maior de informações ao passo que todo o processo tende a ser mais demorado pois envolve tarefas de normalização, segmentação e processamento de todos os dados.

Apesar do método deste estudo não envolver análise de imagens médicas, podemos ter como referência os classificadores utilizados como também suas métricas de performance.

(ZHANG et al., 2015) obteve as seguintes métricas com seus modelos respectivos:

1. kernel polinomial, AUC 92,4%
2. kernel linear, AUC 91,5%
3. função base radial, AUC 86,7%

Os ensaios de (MORADI et al., 2015), utilizando a abordagem de *Random Forest* com o conjunto de dados de (YE; POHL; DAVATZIKOS, 2011), obteve AUC de 55%.

Metodologia

Primeiramente, todo o conjunto de dados foi avaliado para buscar informações não coerentes, atributos nulos ou atributos discrepantes. Dados nulos ou ausentes serão estudados para verificar a viabilidade de aplicar imputação de dados baseado, por exemplo, na média do conjunto de dados. Por outro lado, os dados discrepantes

ou não coerentes deverão ser removidos pois eles podem trazer algum viés e prejudicar no desenvolvimento do modelo.

Na etapa de treinamento foi utilizado, o conjunto de dados foi dividido em 70% para treinamento e 30% para testes. Com o objetivo de maximizar a generalização do modelo de estudo, foi utilizado a técnica de *Cross Validation*.

Este estudo avaliou dois algoritmos de classificação *Support Vector Machines (SVM)* e *Random Forest*. Tais algoritmos foram explorados dada sua capacidade de separação de classes binárias e generalização, e por também ser amplamente utilizada em estudos médicos de classificação de dados (LI; CHUNG; WANG, 2015; EBRAHIMI et al., 2017; TAYEFI et al., 2017; LEBEDEV et al., 2014; AHMAD; MOURSHED; REZGUI, 2017; SHAIKHINA et al., 2017).

Pré-processamento de dados

Antes de iniciar os ensaios com os algoritmos de classificação, foi investigado a qualidade dos dados em busca de dados nulos exibido no código-fonte.

```
nans = lambda df: df[df.isnull().any(axis=1)]
```

Foram encontrados valores nulos apenas para o atributo SES e MMSE.

◆	Subject ID ◆	MRI ID ◆	Group ◆	Visit ◆	MR Delay ◆	M/F ◆	Hand ◆	Age ◆	EDUC ◆	SES ◆	MMSE ◆	CDR ◆	eTIV ◆	nWBV ◆	ASF ◆
2	OAS2_0002	OAS2_0002_MR1	Demented	1	0	M	R	75	12	NaN	23.0	0.5	1678	0.736	1.046
3	OAS2_0002	OAS2_0002_MR2	Demented	2	560	M	R	76	12	NaN	28.0	0.5	1738	0.713	1.010
4	OAS2_0002	OAS2_0002_MR3	Demented	3	1895	M	R	80	12	NaN	22.0	0.5	1698	0.701	1.034
10	OAS2_0007	OAS2_0007_MR1	Demented	1	0	M	R	71	16	NaN	28.0	0.5	1357	0.748	1.293
11	OAS2_0007	OAS2_0007_MR3	Demented	3	518	M	R	73	16	NaN	27.0	1.0	1365	0.727	1.286
12	OAS2_0007	OAS2_0007_MR4	Demented	4	1281	M	R	75	16	NaN	27.0	1.0	1372	0.710	1.279
134	OAS2_0063	OAS2_0063_MR1	Demented	1	0	F	R	80	12	NaN	30.0	0.5	1430	0.737	1.228
135	OAS2_0063	OAS2_0063_MR2	Demented	2	490	F	R	81	12	NaN	27.0	0.5	1453	0.721	1.208
207	OAS2_0099	OAS2_0099_MR1	Demented	1	0	F	R	80	12	NaN	27.0	0.5	1475	0.762	1.190
208	OAS2_0099	OAS2_0099_MR2	Demented	2	807	F	R	83	12	NaN	23.0	0.5	1484	0.750	1.183
237	OAS2_0114	OAS2_0114_MR1	Demented	1	0	F	R	76	12	NaN	27.0	0.5	1316	0.727	1.333
238	OAS2_0114	OAS2_0114_MR2	Demented	2	570	F	R	78	12	NaN	27.0	1.0	1309	0.709	1.341
322	OAS2_0160	OAS2_0160_MR1	Demented	1	0	M	R	76	12	NaN	27.0	0.5	1557	0.705	1.127
323	OAS2_0160	OAS2_0160_MR2	Demented	2	552	M	R	78	12	NaN	29.0	1.0	1569	0.704	1.119
356	OAS2_0181	OAS2_0181_MR1	Demented	1	0	F	R	74	12	NaN	26.0	0.5	1171	0.733	1.499
357	OAS2_0181	OAS2_0181_MR2	Demented	2	539	F	R	75	12	NaN	NaN	1.0	1169	0.742	1.501
358	OAS2_0181	OAS2_0181_MR3	Demented	3	1107	F	R	77	12	NaN	NaN	1.0	1159	0.733	1.515
359	OAS2_0182	OAS2_0182_MR1	Demented	1	0	M	R	73	12	NaN	23.0	0.5	1661	0.698	1.056
360	OAS2_0182	OAS2_0182_MR2	Demented	2	776	M	R	75	12	NaN	20.0	0.5	1654	0.696	1.061

Figura 14: Valores nulos encontrados

Para os atributos onde foram encontrados os valores nulos foi aplicado a técnica de imputação de dados pela média que consiste em calcular a média dos demais sujeitos para completar os dados faltantes.

```
df_long["SES"].fillna(df_long["SES"].mean(), inplace=True)
df_long["MMSE"].fillna(df_long["MMSE"].mean(), inplace=True)
```

De forma igual, foi também investigado a ocorrência de valores iguais a zero.

```
df_long.isin([0]).sum()
```

```
Subject ID      0
MRI ID         0
Group          0
Visit          0
MR Delay       150
M/F            0
Hand           0
Age            0
EDUC           0
SES            0
MMSE           0
CDR            206
eTIV           0
nWBV           0
ASF            0
dtype: int64
```

Neste caso, para os atributos utilizados, a ocorrência de valores iguais a zero são esperados. Para o atributo "CDR" o valor igual a zero indica um sujeito normal. Para o atributo "MR Delay" indica que não houve espera durante o exame, isto implica na qualidade e contraste dos exames de imagem de RM.

Por fim, a investigação de dados discrepantes foi realizado por meio do Z-Score (ou Pontuação Z). O Z-Score é o número de desvios padrão de uma observação a partir da média. Uma pontuação de 0 (zero) significa que a pontuação é igual à média. Uma pontuação Z pode ser positiva ou negativa, indicando se está acima ou abaixo da média e o número de desvios padrão.

```
import numpy as np

def outliers_z_score(ys):
    threshold = 3

    mean_y = np.mean(ys)
    stdev_y = np.std(ys)
    z_scores = [(y - mean_y) / stdev_y for y in ys]
    return np.where(np.abs(z_scores) > threshold)
```

Após aplicarmos a função nos atributos "M/F", "MR Delay", "Age", "EDUC", "SES", "MMSE", "CDR", "eTIV", "nWBV", "ASF". Foram identificados poucos pontos discrepantes, dos quais ainda assim indicam condição dentro da normalidade, seja para a espera para coleta de exame de RM ou níveis mais severos da DA indicados por pontuações elevadas dos testes de cognição como o MMSE e o CDR.

```
Atributo: M/F, Outliers (array([], dtype=int64),):
Atributo: MR Delay, Outliers (array([ 75, 160, 265]),):
Atributo: Age, Outliers (array([], dtype=int64),):
Atributo: EDUC, Outliers (array([], dtype=int64),):
Atributo: SES, Outliers (array([], dtype=int64),):
Atributo: MMSE, Outliers (array([ 26,  89, 100, 101, 172, 173, 251]),):
Atributo: CDR, Outliers (array([184, 251, 330]),):
Atributo: eTIV, Outliers (array([], dtype=int64),):
Atributo: nWBV, Outliers (array([], dtype=int64),):
Atributo: ASF, Outliers (array([], dtype=int64),):
```

Após a normalização dos valores do conjunto de dados, foi aplicado dicionários para cada atributo para posteriormente ser utilizados nos algoritmos de classificação. Os dicionários utilizados visa associar um nome, como "M" de masculino e "F" de feminino para 1 e 0, por exemplo. A seguir é exibido os dicionários utilizados no conjunto de dados, especialmente nos atributos de classe "Group", de sexo "M/F" e mão dominante (canhoto/destro) "Hand".

```
group_map = {"Demented": 1, "Nondemented": 0}
gender_map = {"M": 1, "F": 0}
hand_map = {"R": 1, "L": 0}

df_long['Group'] = df_long['Group'].replace(['Converted'], ['Demented'])
df_long['Group'] = df_long['Group'].map(group_map)
df_long['M/F'] = df_long['M/F'].map(gender_map)
df_long['Hand'] = df_long['Hand'].map(hand_map)
```

Implementação

Este estudo utilizou basicamente dois algoritmos de classificação, o Random Forest e o Support Vector Machine (SVM). Tais algoritmos lidam bem com o cenário do estudo, ao modelar duas classes, o que chamamos de um classificador binário.

O primeiro experimento utilizou os valores padrões de cada algoritmo.
Primeiro para o Random Forest.

```
from sklearn.ensemble import RandomForestClassifier

rf_model = RandomForestClassifier(max_depth=2, random_state=0)
rf_model.fit(X_train, y_train.ravel())

RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
                        max_depth=2, max_features='auto', max_leaf_nodes=None,
                        min_impurity_decrease=0.0, min_impurity_split=None,
                        min_samples_leaf=1, min_samples_split=2,
                        min_weight_fraction_leaf=0.0, n_estimators=10, n_jobs=1,
                        oob_score=False, random_state=0, verbose=0, warm_start=False)
```

Os resultados na matriz de confusão obtido o seguinte.

Classe	Precisão	Sensibilidade	Pontuação F1	Suporte
Não demência	86%	98%	92%	52
Demência	98%	87%	92%	60
Média	93%	92%	92%	112

Tabela 4:
Matriz de confusão para o classificador Random Forest sem otimizações

Do mesmo modo, foi utilizado um classificador padrão, inicial para o SVM, com kernel linear.

```
from sklearn import svm

svm_linear_model = svm.SVC(kernel='linear')
svm_linear_model.fit(X_train, y_train.ravel())
```

```
SVC(C=1.0, cache_size=200, class_weight=None, coef0=0.0,
    decision_function_shape='ovr', degree=3, gamma='auto', kernel='linear',
    max_iter=-1, probability=False, random_state=None, shrinking=True,
    tol=0.001, verbose=False)
```

Apesar de esperarmos resultados melhores, inicialmente o SVM foi ligeiramente pior no indicador de sensibilidade, o mais importante deste estudo.

Classe	Precisão	Sensibilidade	Pontuação F1	Suporte
Não demência	86%	96%	91%	52
Demência	96%	87%	91%	60
Média	92%	91%	91%	112

Tabela 5:
Matriz de confusão para o classificador Support Vector Machine (SVM) sem otimizações

Refinamento

A estratégia adotada para realizar os ajustes finos de cada algoritmo de classificador se baseou em utilizar o GridSearchCV. Este seletor de modelo também traz a validação cruzada que torna o classificador mais robusto ao poder lidar com uma variedade de maior de novos exemplos, pois se baseia na média dos parâmetros ajustados nas validações cruzadas utilizadas.

Foram utilizadas configurações de GridSearchCV visando obter maior pontuação scoring de sensibilidade recall e também de ROC AUC roc_auc. Os resultados com o objetivo de maximizar o ROC AUC foram os que obtiveram melhores resultados.

```
from scipy.stats import randint as sp_randint
from sklearn.model_selection import GridSearchCV

param_grid = {"n_estimators" : [10, 15, 20, 25],
              "max_depth": [2, 4, 6, 8, 12, 14],
              "max_features": ['auto', 9],
              "min_samples_split": [2, 4, 6, 8, 10],
              "min_samples_leaf": [1, 2, 4, 6, 8, 10],
              "bootstrap": [True, False],
              "criterion": ["gini", "entropy"]}

clf_rf = RandomForestClassifier(n_estimators=20)

grid_search = GridSearchCV(clf_rf, param_grid=param_grid, cv=10, scoring="roc_auc")
grid_search.fit(X, y)
report(grid_search.cv_results_)

Model with rank: 1
Mean validation score: 0.921 (std: 0.025)
Parameters: {'bootstrap': False, 'criterion': 'gini', 'max_depth': 12, 'max_features': 'auto',
'min_samples_leaf': 2, 'min_samples_split': 2, 'n_estimators': 25}

Model with rank: 2
Mean validation score: 0.917 (std: 0.031)
Parameters: {'bootstrap': False, 'criterion': 'gini', 'max_depth': 12, 'max_features': 'auto',
```

```
'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 25}
```

Model with rank: 3

Mean validation score: 0.915 (std: 0.020)

Parameters: {'bootstrap': False, 'criterion': 'gini', 'max_depth': 14, 'max_features': 'auto', 'min_samples_leaf': 1, 'min_samples_split': 6, 'n_estimators': 25}

Para o SVM, o grid de parâmetros utilizados para as validações utilizou outros tipos de kernel, além do linear.

```
param_grid_svm = [  
    {'C': [1, 10, 1000], 'kernel': ['linear'], 'shrinking': [True, False],  
     'decision_function_shape': ['ovo', 'ovr', None]},  
    {'C': [1, 10, 1000], 'kernel': ['rbf'], 'degree': [2, 4], 'gamma': [0.1, 0.01, 0.001],  
     'shrinking': [True, False], 'decision_function_shape': ['ovo', 'ovr', None]},  
    {'kernel': ['poly', 'sigmoid'], 'degree': [2, 4], 'coef0': [0, 1, 2, 3], 'shrinking': [True, False],  
     'decision_function_shape': ['ovo', 'ovr', None]}  
]
```

```
from sklearn.model_selection import GridSearchCV
```

```
clf_svm = svm.SVC()
```

```
grid_search = GridSearchCV(clf_svm, param_grid=param_grid_svm, cv=10, scoring="roc_auc")
```

```
grid_search.fit(X, y)
```

```
report(grid_search.cv_results_)
```

Model with rank: 1

Mean validation score: 0.962 (std: 0.018)

Parameters: {'C': 1, 'decision_function_shape': 'ovo', 'degree': 2, 'gamma': 0.1, 'kernel': 'rbf', 'shrinking': True}

Model with rank: 1

Mean validation score: 0.962 (std: 0.018)

Parameters: {'C': 1, 'decision_function_shape': 'ovo', 'degree': 2, 'gamma': 0.1, 'kernel': 'rbf', 'shrinking': False}

Model with rank: 1

Mean validation score: 0.962 (std: 0.018)

Parameters: {'C': 1, 'decision_function_shape': 'ovo', 'degree': 4, 'gamma': 0.1, 'kernel': 'rbf', 'shrinking': True}

Model with rank: 1

Mean validation score: 0.962 (std: 0.018)

Parameters: {'C': 1, 'decision_function_shape': 'ovo', 'degree': 4, 'gamma': 0.1, 'kernel': 'rbf', 'shrinking': False}

Model with rank: 1

Mean validation score: 0.962 (std: 0.018)

Parameters: {'C': 1, 'decision_function_shape': 'ovr', 'degree': 2, 'gamma': 0.1, 'kernel': 'rbf', 'shrinking': True}

Model with rank: 1


```

Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': 'ovr', 'degree': 2, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': False}

Model with rank: 1
Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': 'ovr', 'degree': 4, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': True}

Model with rank: 1
Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': 'ovr', 'degree': 4, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': False}

Model with rank: 1
Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': None, 'degree': 2, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': True}

Model with rank: 1
Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': None, 'degree': 2, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': False}

Model with rank: 1
Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': None, 'degree': 4, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': True}

Model with rank: 1
Mean validation score: 0.962 (std: 0.018)
Parameters: {'C': 1, 'decision_function_shape': None, 'degree': 4, 'gamma': 0.1,
'kernel': 'rbf', 'shrinking': False}

```

Resultados

Modelo de avaliação e validação

Foi aplicado as regularizações propostas somente para o classificador Random Forest.

Para o algoritmo SVM as regularizações propostas não melhoraram a performance do modelo configurado com os parâmetros padrão. Os valores de sensibilidade obtidos foram, para não demência 92% e para demência 32%.

```

from sklearn.ensemble import RandomForestClassifier

rf_model_opt = RandomForestClassifier(bootstrap=False, criterion="gini", max_depth=12,
                                     n_estimators=25, max_features="auto", min_samples_leaf=2,
                                     min_samples_split=2, random_state=0)
rf_model_opt.fit(X_train, y_train.ravel())

```

O algoritmo de classificação Random Forest foi o que se saiu melhor quando comparado com o SVM, por uma diferença mínima. A grosso modo ambos algoritmos tiveram a mesma performance.

Classe	Precisão	Sensibilidade	Pontuação F1	Suporte
Não demência	91%	98%	94%	52
Demência	98%	92%	95%	60
Média	95%	95%	95%	112

Tabela 6:
Matriz de confusão para o classificador Random Forest com otimizações

Além disso avaliamos o parâmetro AUC (Area Under the Curve) de modo a podermos comparar com os estudos de benchmarks citados para avaliar se os resultados encontrados neste estudo estão no mesmo patamar de estudos similares.

```
import numpy as np
from sklearn.metrics import roc_auc_score

roc_auc_score(y_test, rf_model_opt.predict(X_test))
roc_auc_score(y_test, svm_linear_model.predict(X_test))
```

O valor de AUC obtido pelo algoritmo de classificação Random Forest foi de 94,87%. O AUC do classificador SVM foi de 91,41%.

Justificativa

Os resultados obtidos estão nos mesmos patamares de estudos semelhantes, o que demonstra que os algoritmos utilizados desempenharam bem, dado o conjunto de dados utilizado.

A comparação do parâmetro de pontuação AUC com os referência demonstram que o classificador utilizado neste estudo possuiu performance igual ou superior.

Autor	Classificador	AUC
(ZHANG et al., 2015)	SVM com kernel polinomial	92,4%
(ZHANG et al., 2015)	SVM com kernel linear	91,5%
Este estudo	SVM com kernel linear	91,4%
(ZHANG et al., 2015)	SVM com kernel linear	91,5%
(ZHANG et al., 2015)	SVM com função base radial	86,7%
(MORADI et al., 2015)	Random Forest	55%
Este estudo	Random Forest	94,87%

Tabela 7:
Matriz de confusão para o classificador Random Forest com otimizações

Conclusão

Forma livre de visualização

O classificador com melhor desempenho, o Random Forest, elencou os atributos da figura em ordem de importância, a começar pelo gênero, idade, nível de educação. índice sócio-econômico, seguido pelos testes de cognição, e finalmente os parâmetros biológicos.

Estes dados são importantes pois estão coerentes primeiramente com o conjunto de dados, onde não há um equilíbrio de sexo masculino e feminino, a qual não podemos concluir que o gênero é um atributo discriminador

dado o que se conhece da demência. Por outro lado, os demais atributos estão de acordo com os estudos da área, destacando como atributos importantes de determinação da demência, a idade e condições sócio-culturais.

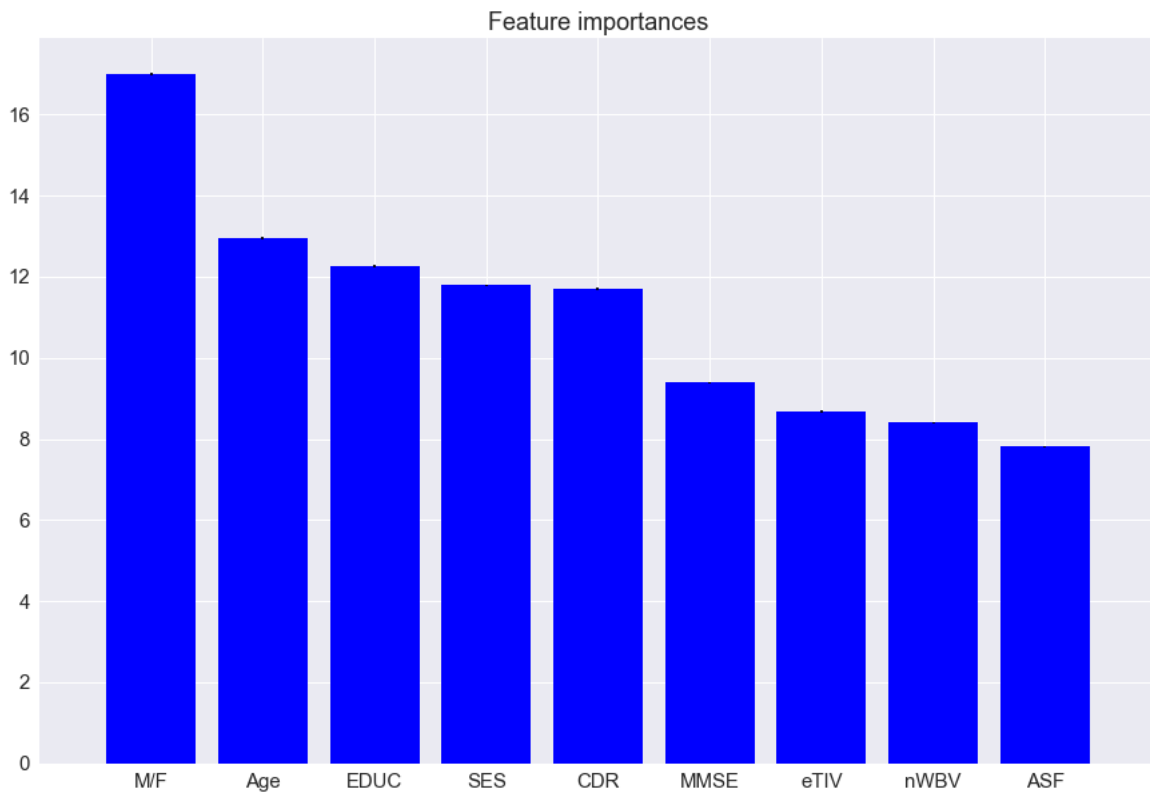


Figura 15: Importância de atributos

Utilizando destes atributos, de forma adequada, o classificador foi capaz de separar as classes com nível relevante de sensibilidade na identificação da demência.

Reflexão

Este estudo comprova o quanto as técnicas tanto avançadas quanto populares de aprendizado de máquina podem impactar positivamente o diagnósticos de certas doenças, podendo em alguns casos realizar predições com antecedência suficiente para que estratégias e cuidados especiais sejam tomados no sentido de ampliar a vida útil dos pacientes e tornar seu tratamento menos agressivo e mais confortável.

Atualmente, iniciativas de dados abertos, para promover a colaboração de diferentes pesquisadores torna achados de estudos como este cada vez mais frequentes nas comunidades dentro e fora da academia. Aliado a organizações como o Kaggle⁶), onde podemos compartilhar nossos algoritmos com especialistas, é possível escalar ainda mais os avanços estudados.

Acredito que a evolução natural deste estudo é agregar mais atributos, aumentar o seguimento dos pacientes, e utilizar os dados de exames posteriores como validadores do modelo e um indicador de antecedência de classificação para os sujeitos que converteram de saudáveis para demência.

Todas as etapas do projeto foram conduzidas conforme o planejamento inicial e não houveram nenhuma grande dificuldade.

⁶<<http://www.kaggle.com>>

Melhorias

O uso de outros tipos de classificadores, de forma exaustiva, poderia trazer luz a outros métodos de classificação diferente dos amplamente utilizados em estudos similares, como por exemplo Naive Bayes Multinomial, métodos ensemble como AdaBoost e XGBoost e até mesmo outras abordagens utilizando Deep Learning. Aliado a facilidade de automatizar tais testes, recorrendo inclusive a computação de nuvem.

O conjunto de dados utilizado ainda é pequeno, o que pode trazer algum tipo de viés aos grupos de estudo. O ideal seria contar com bases maiores, representando, pelo menos 1000 elementos de cada grupo, respeitando a distribuição nos diferentes atributos utilizados como parâmetros de classificação.

Ainda há espaço para incluir outros atributos importantes relacionado a pacientes com demência, tais como exames de líquido cefalorraquidiano (líquor) dada a sua eficácia em medir proteínas causadoras da DA, informações genéticas.

Por fim, um ponto relevante de estudos médicos é o poder de validação do modelo com o seguimento dos pacientes.

Neste conjunto de dados, houve um seguimento pequeno de um grupo de sujeitos que evoluíram do estado de não demência para demência. Esta evolução é extremamente importante pois nos traz valiosas observações da importância do peso dos atributos utilizados e se torna prova real do poder de classificação dos modelos utilizados.

Glossário

3MS *Modified Mini-Mental State*. 12

BO Bandas Oligoclonais. 12

BOMC *Blessed Orientation Memory Concentration Test*. 12

CDR *Clinical Dementia Rating*. 12

CJD *Creutzfeldt-Jakob Disease*. 12

DSM-IV *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition*. 12

DV Demência Vascular. 12

EEG Eletroencefalograma. 12

FN Falso Negativo. 12

IgG *Imunoglobulina G*. 12

MMSE *Mini-Mental State Examination*. 12

NINCDS-ARDA *National Institute of Neurologic, Communicative Disorders and Stroke/Alzheimers Disease and Related Disorders Association*. 12

PET *Positron Emission Tomography.* 12

pib Produto Interno Bruto. 12

SPECT *Single Photon Emission Computed Tomography.* 12

STMS *Short Test of Mental Status.* 12

VP Verdadeiro Positivo. 12

Referências

- AHMAD, M. W.; MOURSHED, M.; REZGUI, Y. Trees vs neurons: Comparison between random forest and ann for high-resolution prediction of building energy consumption. *Energy and Buildings*, v. 147, p. 77 – 89, 2017. ISSN 0378-7788. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378778816313937>>.
- BEAL, M. F.; LANG, A. E.; LUDOLPH, A. C. *Neurodegenerative Diseases*. [S.l.]: Cambridge University Press, 2005.
- BLESSED, G.; TOMLINSON, B. E.; ROTH, M. The association between quantitative measures of dementia and of senile changes in the cerebral gray matter of elderly subjects. *The British Journal of Psychiatry*, p. 797–811, 1968.
- CHUNG, J. A.; L., C. J. Neurobehavioral and neuropsychiatric symptoms in alzheimer's disease: Characteristics and treatment. *Neurologic Clinics*, p. 829–46, 2000.
- CLAUSS, J. J.; HASKAMP, F. van; BRETELER, M. M. The diagnostic value of spect with tc 99m hmpao in alzheimer's disease: a population-based study. *Neurology*, p. 454–61, 1994.
- CORDER, E. H.; SAUNDERS, A. M.; STRITTMATTER, W. J. Gene dose of apolipoprotein e type 4 allele and the risk of alzheimer's disease in late-onset families. *Science*, p. 921–3, 1993.
- D., M. In-office mental status testing: a practical guide. *Geriatrics*, 1991.
- EBRAHIMI, M. et al. Vision-based pest detection based on svm classification method. *Computers and Electronics in Agriculture*, v. 137, p. 52 – 58, 2017. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016816991631136X>>.
- ERTEKIN-TANER, N.; GRAFF-RADFORD, N.; YOUNKIN, L. Linkage of plasma abeta42 to a quantitative locus on chromosome 10 in late-onset alzheimer's disease pedigrees. *Science*, p. 2303–4, 2000.
- FERNANDES, M. P. et al. Avanços em visão computacional. *Alzheimer's Dementia: Translational Research Clinical Interventions*, Omnipax, v. 1, n. 1, p. 1 – 24, 2012.
- FOLSTEIN, M. F.; FOLSTEIN, S. E.; MCHUGH, P. R. Mini-mental state. a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, p. 189–98, 1975.
- FORLENZA, O. V.; ALMEIDA, J. R. *Fisiopatologia dos Transtornos Psiquiátricos*. [S.l.: s.n.], 2006.
- FORSTL, H.; LEVY, R. On certain peculiar diseases of old age. *History of Psychiatry*, p. 71–101, 1991.
- FOX, N. C. et al. Correlation between rates of brain atrophy and cognitive decline in ad. *Neurology*, p. 1687–9, 1999.
- FOX, N. C.; WARRINGTON, E. K.; FREEBOROUGH, P. A. Presymptomatic hippocampal atrophy in alzheimer's disease: A longitudinal study. *Brain*, p. 2001–7, 1996.
- FREMPONG NIMO NICHOLAS, M. A. B. N. K. Decision tree as a predictive modeling tool for auto insurance claims. *International Journal of Statistics and Applications*, 2017.
- GALASKO, D.; HANSEN, L. A.; KATZMAN, R. Clinical neuropathological correlations in alzheimer's disease and related dementias. *Archives of Neurology*, p. 888–95A, 1994.
- GEORGE-HYSLOP, P. H. S. et al. Two homologous genes causing early-onset familial alzheimer's disease. *Cold Spring Harbor Symposia on Quantitative Biology*, p. 559–64, 1996.
- JACK, C. R. J.; PETERSEN, R. C.; XU, Y. C. Medial temporal atrophy on mri in normal aging and very mild alzheimer's disease. *Neurology*, p. 786–94, 1997.

- JACK, C. R. J.; PETERSEN, R. C.; XU, Y. C. Prediction of ad with mri-based hippocampal volume in mild cognitive impairment. *Neurology*, p. 1397–403, 1999.
- KATZMAN, R. et al. Validation of a short orientation-memory concentration test of cognitive impairment. *American Journal of Psychiatry*, p. 734–9, 1983.
- KLOPPEL, S. et al. Automatic classification of mr scans in alzheimer's disease. *Brain*, 2008.
- KLUNEMANN, H. H. et al. Alzheimer's second patient: Johann f. and his family. *Annals of Neurology*, p. 530–3, 2002.
- KNOPMAN, D. S.; DEKOSKY, S. T.; CUMMINGS, J. L. Practice parameter: Diagnosis of dementia (an evidence-based review). report of the quality standards subcommittee of the american academy of neurology. *Neurology*, p. 1143–53, 2001.
- KOKMEN, E. et al. The short test of mental status: Correlations with standardized psychometric testing. *Archives of Neurology*, p. 725–8, 1991.
- LEBEDEV, A. et al. Random forest ensembles for detection and prediction of alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*, v. 6, p. 115 – 125, 2014. ISSN 2213-1582. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2213158214001326>>.
- LI, H.; CHUNG, F. lai; WANG, S. A svm based classification method for homogeneous data. *Applied Soft Computing*, v. 36, p. 228 – 235, 2015. ISSN 1568-4946. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1568494615004676>>.
- LIN, K. A. et al. Marked gender differences in progression of mild cognitive impairment over 8 years. *Alzheimer's Dementia: Translational Research Clinical Interventions*, v. 1, n. 2, p. 103 – 110, 2015. ISSN 2352-8737. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2352873715000190>>.
- MCKHANN, G. et al. Clinical diagnosis of alzheimer's disease: Report of the nincds-adrda work group under the auspices of department of health and human services task force on alzheimer's disease. *Neurology*, p. 939–44, 1984.
- MORADI, E. et al. Machine learning framework for early mri-based alzheimers conversion prediction in mci subjects. *NeuroImage, Elsevier*, v. 104, p. 398–412, 1 2015. ISSN 1053-8119.
- MORRIS, J. C. The clinical dementia rating (cdr): Current version and scoring rules. *Neurology*, p. 2412–14, 1993.
- MUNHOZ, D.; FELDMAN, H. Causes of alzheimer's disease. *Canadian Medical Association Journal*, p. 65–72, 2002.
- SHAIKHINA, T. et al. Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. *Biomedical Signal Processing and Control*, 2017. ISSN 1746-8094. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1746809417300204>>.
- SHERRINGTON, R.; ROGAEV, E. I.; LIANG, Y. Cloning of a gene bearing mis-sense mutations in early-onset familial alzheimer's disease. *Nature*, p. 754–60, 1995.
- SMALL, G.; RABINS, P.; BARRY, P. Diagnosis and treatment of alzheimer's disease and related disorders. consensus statement of the american association for geriatric psychiatry, the alzheimer's association, and the american geriatrics society. *Journal of the American Medical Association*, p. 1363–71, 1997.
- STELZMANN, R. A.; SCHNITZLEIN, H. N.; MURTAGH, F. R. An english translation of alzheimer's 1907 paper 'über eine eigenartige erkankung der hirnrinde'. *Clinical Anatomy*, p. 429–31, 1995.

- SUNDHARI, S. S. A knowledge discovery using decision tree by gini coefficient. In: *2011 International Conference on Business, Engineering and Industrial Applications*. [S.l.: s.n.], 2011. p. 232–235.
- TANNER, M. S. L. et al. Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness. *PLoS Neglected Tropical Diseases*, 2008.
- TAYEFI, M. et al. The application of a decision tree to establish the parameters associated with hypertension. *Computer Methods and Programs in Biomedicine*, v. 139, p. 83 – 91, 2017. ISSN 0169-2607. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169260716303777>>.
- TEKIN, S.; CUMMINGS, J. L. Depression in dementia. *The Neurologist*, p. 252–9, 2001.
- TENG, E. L.; CHUI, H. C. The modified mini-mental state (3ms) examination. *Journal of Clinical Psychiatry*, p. 314–18, 1987.
- VANGOOL, W. A. et al. Diagnosing alzheimer's disease in elderly, mildly demented patients: the impact of routine single photon emission computed tomography. *Journal of Neurology*, p. 401–5, 1995.
- YE, D. H.; POHL, K.; DAVATZIKOS, C. Semi-supervised pattern classification: application to structural mri of alzheimer's disease. *Pattern Recognition in NeuroImaging (PRNI)*, International Workshop on. IEEE, v. 1, p. 1–4, 1 2011. ISSN 1053-8119.
- ZHANG, Y. et al. Detection of subjects and brain regions related to alzheimer's disease using 3d mri scans based on eigenbrain and machine learning. *Frontiers in Computational Neuroscience*, v. 9, p. 66, 2015. ISSN 1662-5188. Disponível em: <<https://www.frontiersin.org/article/10.3389/fncom.2015.00066>>.
- ZUBENKO, G. S.; ZUBENKO, W. N.; MCPHERSON, S. A collaborative study of the emergence and clinical features of the major depressive syndrome of alzheimer's disease. *American Journal of Psychiatry*, p. 857–66, 2003.
- ÇAMLICA, Z.; TIZHOOSH, H. R.; KHALVATI, F. Medical image classification via svm using lbp features from saliency-based folded data. In: *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*. [S.l.: s.n.], 2015. p. 128–132.