



DOI: 10.5335/rbca.v11i2.9037 Vol. 11, No 2, pp. 31-41

Homepage: seer.upf.br/index.php/rbca/index

ARTIGO ORIGINAL

CINTIA 2: uma hierarquia de redes neurais artificiais binárias para classificação inteligente de supernovas

Francisca Joamila Brito do Nascimento¹, Luis Ricardo Arantes Filho¹ and Lamartine Nogueira Frutuoso Guimarães^{1,2}

¹Instituto Nacional de Pesquisas Espaciais and ²Instituto de Estudos Avançados

Recebido: 12/01/2019. Revisado: 23/04/2019. Aceito: 07/05/2019.

Resumo

Supernovas são eventos catastróficos no qual algumas estrelas explodem. A classificação de supernovas é feita por especialistas por meio da análise dos espectros de luz que apresentam linhas de absorção e emissão em determinadas regiões do comprimento de onda. Os espectros de luz das supernovas apresentam padrões que podem ser usados em algoritmos de aprendizagem de máquina possibilitando assim a classificação automática e inteligente das supernovas. A classificação automática é essencial para o processamento de grande quantidade de dados em equipamentos instalados em lugares remotos, onde não é sempre possível a presença de um especialista. O objetivo deste trabalho é apresentar a CINTIA 2, aprimoramento do Classificador Inteligente de supernovas do tipo Ia, que usa uma hierarquia de redes neurais binárias do tipo Perceptron para classificar as supernovas nos tipos Ia, Ib, Ic e II. Apresentamos a arquitetura da CINTIA 2 e a ferramenta daí proveniente, desenvolvida nas linguagens de programação Python e C++. Os resultados obtidos apresentam ótimo desempenho, principalmente na classificação dos tipos Ia e II. Uma comparação com trabalhos encontrados na literatura mostra que a CINTIA 2 é superior em quantidade e diversidade de dados e alcança índices de classificação equiparáveis aos demais classificadores.

Palavras-Chave: classificador inteligente; hierarquia; redes neurais artificiais; supernovas

Abstract

Supernovae are catastrophic events of stars explosion. The classification of supernovas is done by specialists by means of the analysis of the light spectrum that have lines of absorption and emission in certain regions of the wavelength. The supernovae light spectrua present patterns that can be used in machine learning algorithms, thus enabling automatic and intelligent classification of supernovae. Automatic classification is essential for the processing of large amounts of data in equipment installed in remote locations, where it is not always possible to have a specialist. The aim of this work is to present CINTIA 2, an enhancement of the Intelligent Classifier of Type Ia supernovae (CIntIa), which uses a hierarchy of binary neural networks of the Perceptron kind to classify supernovas in types Ia, Ib, Ic and II. We present the architecture of CINTIA 2 and the tool derived from it, developed in the programming languages Python and C ++. The results obtained presented excellent performance, mainly in the classification of types Ia and II. A comparison with works found in the literature shows that CINTIA 2 is superior in quantity and diversity of data and reaches classification indexes comparable to the other classifiers.

Key words: artificial neural networks; hierarchy; intelligent classifier; supernovae

^{*}francisca.nascimento@inpe.br; luisricardoengcomp@gmail.com; guimarae@ieav.cta.br

1 Introdução

Supernovas (SNs) são eventos catastróficos no qual estrelas muito massivas, em sua maioria, explodem. Essa explosão emite brilho muito intenso, segundo Damineli and Steiner (2010), as grandes estrelas, ao sucumbir, superam galáxias inteiras em brilho. Seus clarões podem ser vistos por toda a extensão do Universo por alguns dias. Além das estrelas muito massivas (>10⊙), SNs também podem ser causadas pela explosão de anãs brancas que atraíram para si massa de sua companheira em um sistema binário. As SNs fruto da explosão de estrelas muito massivas dizemos que são de colapso de núcleo, enquanto as outras denominamos de SNs termonucleares.

As SNs termonucleares são classificadas como Ia (SNIa) e de acordo com Carroll and Ostlie (2007) contam com uma forte presença de linhas espectrais de Silício (Si). As SNIa são muito importantes para os estudos de Cosmologia porque podem ser usadas como velas-padrão. O melhor exemplo disso foi o estudo que provou a expansão acelerada do Universo, rendendo o prêmio Nobel de Física de 2011 para Saul Perlmutter Perlmutter et al. (1999), Adam G. Riess e Brian P. Schmidt Riess et al. (1998).

A classificação de SNs é feita por especialistas por meio da análise dos espectros de luz que apresentam linhas de absorção e emissão em determinadas regiões do comprimento de onda, evidenciando a presença de alguns elementos químicos. A Figura 1, publicada originalmente por Turatto (2003), mostra quais elementos químicos são considerados para efetuar a classificação. As SNIa tem presença marcante de Si, que se evidencia, principalmente, nas linhas de absorção em torno de λ 6135 $\hbox{\AA}$ e Enxofre (S). Enquanto as SNs do tipo II (SNII) apresentam Hidrogênio (H), ausente nas SNs do tipo I e as SNs Ib (SNIb) tem predominância de Hélio (He). As SNs Ic (SNIc) são classificadas considerando a ausência dos quatro elementos químicos supracitados.

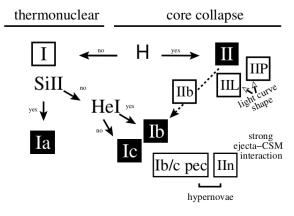


Figura 1: Esquema de classificação das SNs de acordo com os elementos químicos presentes no espectro.

Os espectros de luz das SNs apresentam padrões que podem ser usados em algoritmos de aprendizagem de máquina possibilitando assim a classificação automática das SNs. O emprego de classificação automática de SNs, abordado em alguns trabalhos encon-

trados na literatura, é justificável pelo aumento de telescópios instalados ao redor do globo terrestre e a quantidade de dados produzida por eles diariamente. Esse será o caso do Projeto Kulun Dark Universe Survey Telescope (KDUST), telescópio que está sendo instalado na Antártida pelo governo chinês com previsão de conclusão para o ano de 2020 Zhu et al. (2014). Estratégias de classificação automática são essenciais para o processamento de grande quantidade de dados em equipamentos instalados em lugares remotos, onde não é sempre possível a presença de um especialista em SNs.

No contexto de classificadores automáticos, a CIntIa (Classificador Inteligente de supernovas do tipo Ia) Módolo et al. (2016) foi desenvolvida para efetuar a classificação de SNs utilizando redes neurais artificiais. A primeira versão, que utilizou dados provenientes do acervo do Harvard-Smithsonian Center for Astrophysics (Cfa) (Blondin et al. (2011), Blondin et al. (2012), Matheson et al. (2008) e Modjaz et al. (2014)) apresenta um bom desempenho na classificação de SNIa. A classificação de SNIb, SNIc e SNII é difícil mensurar dado a pouca quantidade de dados disponíveis nos testes. A utilização de apenas uma base de dados mais a introdução de algumas restrições contribuiu para a pequena quantidade de dados. Restrições essas que incluem a utilização apenas de espectros com fase espectral entre -3 e +7 dias. A fase espectral corresponde ao dia em que o espectro foi obtido, tendo como referência o dia em que a SN alcançou seu brilho máximo que é considerada a fase o (zero). E estratégia de filtragem simples que não foi suficiente para eliminar ruídos de determinados espectros, que assim foram descartados do processo.

O objetivo deste trabalho é apresentar a segunda versão da CIntIa, denominada CINTIA 2, que usa uma hierarquia de redes neurais artificiais binárias do tipo Perceptron para classificar as SNs nos 4 tipos introduzidos acima. Cada rede neural binária é um módulo que determina se um padrão pertence ou não a uma classe, estratégia inspirada por Anad et al. (1995). A quantidade de espectros foi incrementada em mais de 1300% (de 649 para 9156), possibilitando assim uma avaliação mais robusta para os 4 tipos (Ia, Ib, Ic e II) do que a proposta por Módolo et al. (2016). A arquitetura criada para efetivar a hierarquia das redes neurais possibilita uma classificação não-ambígua de cada espectro, o que não havia sido alcançado pela CIntIa em sua primeira versão. Essa arquitetura foi apresentada em uma versão preliminar em do Nascimento et al. (2019), as mudanças abordadas aqui decorrem de novos testes e aumentam a qualidade do classificador. A arquitetura foi implementada em uma ferramenta desenvolvida nas linguagens Python e C++, permitindo a sua utilização por qualquer pessoa interessada.

Este artigo está organizado da seguinte maneira. A Seção 2 apresenta conceitos de redes neurais artificiais, com ênfase no Perceptron e nas redes neurais binárias. A seção seguinte trata dos dados utilizados no desenvolvimento da CINTIA 2, sua origem, limpeza e pré-processamento. A Seção 3 é uma revisão do estado da arte de classificadores automáticos e /ou inteligentes de SNs. A Seção 4 discorre sobre os aspectos importantes dos treinamentos das redes neurais. A Seção 5 apresenta os resultados e discussão daí decorrentes. Finalmente, na seção seguinte,

é feita uma síntese do que foi apresentado no artigo e enunciamos, brevemente, alguns trabalhos futuros pertinentes.

Background

Uma Rede Neural Artificial (RNA) é uma rede que conecta unidades de processamento simples, os neurônios (Figura 2), e tem a capacidade de aprendizado por meio de exemplos, que denominamos aprendizagem supervisionada. Segundo (HAYKIN, 2001), os elementos que compõem um neurônio são os seguintes: peso, somador, bias e função de ativação.

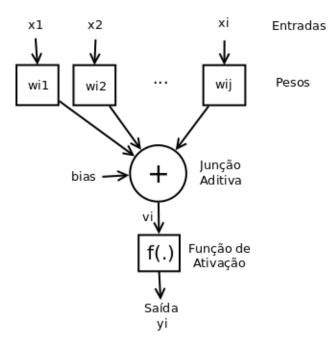


Figura 2: Modelo de um neurônio artificial. Adaptado de Haykin (2001).

O modelo para as RNAs deriva do modelo biológico de funcionamento do cérebro que é composto por bilhões de neurônios. "Um neurônio é uma célula formada por três seções com funções específicas e complementares: corpo, dendritos e axônio. Os dendritos captam os estímulos recebidos em um determinado período de tempo e os transmitem ao corpo do neurônio, onde são processados." Ferneda (2006). Os neurônios transmitem estímulos uns aos outros através das sinapses e assim a informação é processada. De acordo com a argumentação de Haykin (2001), as Equações 1 e 2 descrevem o comportamento matemático do neurônio artificial representado na Figura

$$v_i = \sum_{k=1}^m w_{ik} x_i \tag{1}$$

$$y_i = f(v_i + b_i) \tag{2}$$

No que concerne à função de ativação, as funções

mais usadas são a função limiar, a função linear, a função sigmóide e a função tangente hiperbólica. A função de ativação deve ser escolhida de acordo com a natureza do problema de aprendizagem tratado. A conexão entre os neurônios da RNA pode ser feita em apenas uma camada, na qual as entradas são multiplicadas pelos pesos, passam pelo somador e logo em seguida pela função de ativação. Ou a RNA pode ter mais de uma camada, assim a saída de uma camada é usada como entrada para os neurônios da camada posterior, sendo propagada até a saída.

Há uma grande quantidade de arquiteturas de RNA disponíveis na literatura, cada uma delas usa uma estratégia diferente, na forma de algoritmo, para engendrar o aprendizado. Uma das mais simples e populares é o Multi Layer Perceptron (MLP), que será abordada em seguida.

2.1 Multi Layer Perceptron (MLP)

Uma rede MLP consiste de um conjunto de unidades que constituem a camada de entrada, uma ou mais camadas ocultas e uma camada de saída, como descreve (HAYKIN, 2001). A entrada se propaga camada por camada até chegar a saída. O aprendizado no MLP é comumente realizado com o algoritmo de backpropagation que é baseado na regra de aprendizagem por correção de erro. O aprendizado por correção de erro funciona basicamente aquirindo o sinal de erro da saída e a partir dele os pesos são corrigidos até chegar em um valor aceitável. "A abordagem adotada pelo algoritmo [retropropagação] consiste em iniciar na camada de saída e propagar o erro retroativamente através das camadas ocultas." Luger (2013). O treinamento com retropropagação usa a regra delta generalizada. O cálculo do fator delta de cada neurônio depende se o neurônio está na camada de saída ou em uma camada intermediária.

Resumindo, o algoritmo executa duas etapas principais. O primeiro passo é a propagação para a frente, durante a qual os pesos não são alterados e o sinal de erro é calculado ao fim da propagação. No segundo passo, os deltas derivados do sinal de erro são propagados para trás e os pesos são atualizados. Cada vez que esses dois passos são executados se dá uma sessão de treinamento, também chamada de época de treinamento. São executadas quantas épocas forem necessárias até que a rede tenha aprendido suficientemente de acordo com algum parâmetro chamado condição de parada.

2.2 RNAs binárias

De acordo com Oong and Isa (2012), em reconhecimento de padrões, usar um grupo de redes neurais para resolver um problema é mais efetivo do que usar uma rede neural única. A estratégia conhecida como "One Against All" é usada para solucionar problemas chamados de multiclasses. Um problema multiclasses é aquele em que um objeto pode ser classificado apenas como uma classe, no entanto há mais de duas opções de classe.

No domínio dos MLPs, muitos problemas multiclasses se tornam complexos e a separação das classes não é possível, logo a aprendizagem da rede é comprometida. Portanto, muitos pesquisadores optam por

decompor o problema em subproblemas mais simples. Assim, segundo Mencia and Furnkranz (2008), um conjunto de treino com K possíveis classes é decomposto em K conjuntos binários de teste que são usados para treinar K classificadores binários. Para cada rede neural binária uma classe é considerada positiva e deve apresentar a saída 1, enquanto todas as outras classes são negativas e devem ter como saída o (as saídas 1 e o foram adotadas neste trabalho, mas são descritas com outros valores em outros artigos).

Deve ser considerada ainda a decisão final sobre a qual classe pertence o objeto a ser classificado. Após a classificação feita por cada RNA binária, uma estratégia deve ser adotada para escolher a solução mais apropriada. Mencia and Furnkranz (2008) mencionam os métodos de ranking de classes, que pode ser aleatório ou deliberadamente escolhido. No ranking de classes, cada classe recebe um grau de precedência, então a saída cuja classe tem maior precedência é considerada como a classificação final. Neste trabalho, o método de escolha da classe é feito por uma hierarquia sugerida de acordo com os testes empíricos dos treinamentos efetuados em cada rede neural binária. Ou seja, é praticado um ranking de classes no qual as precedências são escolhidas baseadas nos resultados dos testes efetuados em cada rede neural após os treinamentos.

Classificadores automáticos de SNs

Foram selecionados 6 trabalhos que apresentam classificadores de SNs que usam dados espectroscópicos (espectros de luz), para abordagem nesta seção. Diferentemente da fotometria que usa as curvas de luz (perfil temporal da intensidade do brilho), a classificação por espectroscopia costuma ser imediata porque a análise pode ser feita com apenas um espectro. Essa característica é bastante vantajosa para a aplicação em sistemas automáticos operando em conjunto com telescópios.

3.1 Supernova Identification (SNID)

O SNID é uma ferramenta desenvolvida por Blondin and Tonry (2007) com a finalidade de determinar idade, redshift e classificar SNs. A identificação das SNs, que resulta em sua classificação é feita usando técnicas estatísticas. O espectro que se deseja classificar é correlacionado com outros espectros previamente identificados, o espectro recebe a mesma classificação daquele com quem mais se correlacionou. Foram usados 879 espectros de 65 SNIa, 322 de 19 SNIb/c e 353 de 10 SNII no desenvolvimento do classificador. Os espectros foram previamente pré-processados para serem submetidos à correlação com fluxo normalizado entre o e 1 e filtrados. Os autores consideram que o classificador alcança algum sucesso, mas os resultados não estão explícitos no trabalho.

3.2 Generic Classification Tool (GELATO)

O GELATO, apresentado por Harutyunyan (2008), é um classificador automático de SNs que pretende mitigar a subjetividade das classificações feitas por

especialistas humanos. A ferramenta usa um método matemático para comparar novos espectros com outros previamente classificados de um banco de dados. Primeiramente os espectros são pré-processados com correção do redshift, suavização, reamostragem e a divisão em 11 intervalos do comprimento de onda. A comparação entre os espectros é feita em cada um desses intervalos. A ferramenta computa para cada intervalo a distância relativa entre o novo espectro e todos os espectros do banco de dados. Por fim, calcula-se a média das distâncias relativas, o menor valor se refere ao espectro mais parecido, portanto os espectros recebem a mesma classificação.

3.3 Classificador Inteligente de Supernovas do tipo Ia (CIntIa)

A CIntIa é um classificador inteligente de SNs dos tipos Ia, Ib, Ic e II e foi desenvolvida por Módolo et al. (2016). A ferramenta usa 4 RNAs para reconhecer cada um dos tipos principais de SNs e apresenta ótimos resultados na classificação das SNIa. A CIntia usa 559 espectros de 192 SNIa, 33 espectros de 12 SNIb, 44 espectros de 12 SNIc e 13 espectros de 5 SNII totalizando 649 espectros. Todos os espectros selecionados estão entre 3 dias antes (-3) e 7 dias depois (+7) do brilho máximo das SNs. Os espectros passam por um pré-processamento antes de serem submetidos à RNA, que inclui: correção do redshift, suavização, interpolação, e normalização do fluxo. A etapa seguinte é extrair intervalos do espectro que são usados como entradas nas RNAs.

Nessa versão da CIntIa, o mesmo espectro pode ser classificado de mais de uma maneira diferente. A não integração das quatro RNAs permite a a ambiguidade na classificação, essa chance foi eliminada com a construção da CINTIA 2, que usa a hierarquia para garantir que a classificação seja única. Uma versão anterior da hierarquia que propomos aqui foi apresentada em do Nascimento et al. (2019), porém novos testes realizados permitiram implementar mudanças que melhoram o índice de acertos do classificador.

Dimensionality Reduction and Clustering for Unsupervised Learning in Astronomy (DRACULA)

O DRACULA é um sistema desenvolvido por Sasdelli et al. (2016) para identificar subtipos do tipo Ia. Para isso o DRACULA primeiro reduz a dimensionalidade dos dados e em seguida usa aprendizado nãosupervisionado para realizar a classificação. Foram usados 3677 espectros do tipo Ia na fase de redução de dimensionalidade que foi realizada com a técnica de Deep Learning (DL), que gerou o conjunto de entradas. Em seguida, dos 3677 espectros foram selecionados apenas 486 que foram auferidos entre as fases espectrais -3 e +3 dias. Os 486 espectros selecionados após a redução da dimensionalidade foram submetidos ao algoritmo de aprendizado não-supervisionado K-Médias. O K-Médias divide as amostras em classes de acordo com semelhanças que elas possuam entre si sem que seja fornecido nenhum exemplo.

3.5 Quantitative Classification of Type I Supernovae (Quantitative Classification)

A proposta de Sun and Gal-Yam (2017) é a classificação quantitativa dos subtipos do tipo I de SNs. Na classificação quantitativa, um tipo é identificado de acordo com características que ele possui e não pelas características que ele não possui. O método proposto para classificar quantitativamente é medir a profundidade de linhas de absorção dos elementos Si II (λ 6150 Å) e O I (λ 7774 Å). Foram usados 146 espectros do tipo Ia, 12 do tipo Ib, 19 do tipo Ic e 4 do tipo Ib/c. Os espectros passaram por pré-processamento, que consistiu em: correção do redshift e suavização com o filtro de Savitzky-Golay. A segunda etapa no tratamento dos dados foi encontrar o pseudo-contínuo que delimita as regiões de absorção a ser analisadas, esta etapa foi realizada manualmente por apenas um dos autores. A região foi encontrada e em seguida a profundidade da região foi calculada. As profundidades medidas resultaram em um critério de classificação desenvolvido pelo autores e usado para definir a que tipo pertence cada espectro.

3.6 Sistema Fuzzy Avaliador de Supernovas (SUZAN)

A SUZAN, desenvolvida por Arantes Filho (2018) é um sistema classificador de SNs em seus tipos principais Ia, Ib, Ic e II. O sistema usa Lógica Fuzzy. No total foram usados 3697 espectros de 588 SNs diferentes, dos quais 3082 espectros são de SNIa, 217 do tipo Ib, 282 do tipo Ic e 116 do tipo II. Os espectros passaram pela etapa de pré-processamento que consistiu na Dupla-Filtragem usando o filtro de Savitzky-Golay, ou seja, os espectros foram filtrados duas vezes seguidas e depois o redshift foi ajustado. A SUZAN extrai características dos espectros para em seguida submetê-los ao crivo das regras nebulosas criadas com conhecimento especialista. O resultados são excelentes para a fase espectral de brilho máximo das SNIa, para a qual a modelagem das regras foi focada.

Dados

Os espectros de luz das SNs foram obtidos no repositório aberto The Open Supernova Catalog (Gillochon et al. (2017)), mantido por James Guillochon e Jerod Parent, pesquisadores do CfA. O acervo é uma coletânea que compreende dados de 17 bases de espectros mais contribuições individuais. Nele, estão disponíveis espectros de SNs dos tipos Ia, Ib, Ic, II e os outros tipos que são classificados por fotometria. Foram baixados inicialmente todos os espectros classificados como dos tipos Ia, Ib, Ic e II, excetuando seus subtipos. Cada espectro é um arquivo (.csv) com duas colunas, a primeira é o comprimento de onda e a segunda o fluxo de luminosidade observado ali. O download foi efetuado no dia 30 de abril de 2018, totalizando 12902 espectros de 4883 SNs. As referências das SNs utilizadas não serão listadas aqui, mas podem ser consultadas na aba "Bibliography" do The Open Supernova Catalog¹.

4.1 Limpeza dos dados

A limpeza dos dados inicialmente baixados foi necessária para eliminar todos os arquivos vazios, arquivos com todos os valores de comprimento de onda iguais (formando apenas uma linha reta) e arquivos sem informação sobre a idade do espectro. A idade do espectro se refere a quando o espectro foi coletado em relação ao dia do brilho máximo da SN (fase espectral), o valor referência é o para o dia em que a explosão estava mais luminosa. A informação sobre a fase espectral é importante porque o padrão dos espectros sofre alterações no decorrer do tempo. Após a limpeza, a quantidade de SNs permaneceu a mesma, porém os espectros foram reduzidos para 12516.

4.2 Seleção de intervalo de comprimento de onda

A quantidade de espectros efetivamente utilizada nas etapas de treinamento e teste do classificador é menor que o total resultante da limpeza de dados. Após a etapa de limpeza dos dados foi aplicado um critério de seleção de espectros aptos para a classificação que consiste em excluir aqueles que não possuem intervalo de comprimento de onda entre 4000Å e 7000Å. A escolha desse intervalo decorre da observação das linhas espectrais importantes para a avaliação dos espectros, as mesmas usadas pelos especialistas, as quais estão dentro da faixa de luz visível. A Figura 3 apresenta a quantidade de dados remanescente após a aplicação desse critério de seleção, em comparação com os baixados inicialmente e os que permaneceram após a limpeza dos dados.

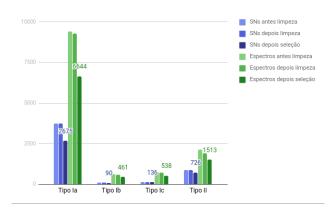


Figura 3: Quantidade de dados baixados, após a limpeza dos dados após a aplicação do critério de seleção intervalo do comprimento de onda.

4.3 Ajuste do Redshift

O redshift é o deslocamento em direção a comprimentos de onda longos (vermelhos), como define Britannica (2018). O ajuste do redshift é recomendado para que possamos analisar o espectro como se os gases da explosão correspondente estivessem em repouso. Dessa forma, as linhas espectrais do evento podem ser comparadas com as linhas medidas em

https://sne.space/bibliography/

laboratório. A correção é feita com a Equação 3.

$$\lambda_0 = \frac{\lambda}{Z+1} \tag{3}$$

Onde:

- λ_0 : comprimento de onda do objeto em repouso;
- λ : comprimento de onda observado;
- · z: redshift.

4.4 Dupla-Filtragem Savitzy-Golay

Faz-se necessário filtrar os espectros para que as redes neurais não tenham a capacidade de aprendizagem comprometida pelos ruídos do sinal obtido na explosão. A estratégia de filtragem utilizada neste trabalho foi aplicada por Arantes Filho (2018) no desenvolvimento do Sistema Fuzzy Avaliador de Supernovas, a SUZAN.

A estratégia de filtragem é chamada de Dupla-Filtragem, pois os espectros passam pelo filtro de Savitzky-Golay (Savitzky and Golay (1964)) duas vezes consecutivas. Os parâmetros usados em cada filtragem são: tamanho da janela de pontos = 71 e grau do polinômio = 9. Na Dupla-Filtragem, inicialmente os espectros são reamostrados em 1000 pontos cada, por uma interpolação linear simples que começa no ponto inicial e termina no ponto final de cada espectro. Em seguida, cada espectro tem seu fluxo normalizado em um vetor de magnitude 1. Então, os espectros são filtrados duas vezes a fim de reduzir ruídos e remover inconsistências.

Treinamento das RNAs

O processo de treinar RNAs consiste em selecionar as entradas pertinentes ao problema, escolher algumas configurações e topologias e testá-las. O teste do desempenho de cada rede neural avalia se ela é capaz de generalizar o aprendizado quando confrontada com exemplos que não foram utilizados durante o treinamento. O treinamento foi efetuado com 60% dos dados, enquanto foram realizados dois testes, cada qual com um conjunto de 20% dos espectros.

A estratégia One Against All, onde treinamos uma rede neural binária para cada classe, resultou na necessidade de treinar 4 RNAs. O treinamento de cada rede foi feito de forma independente e, apesar de não ter sido paralelo é possível que se adote esse procedimento nessa metodologia.

5.1 Extração das entradas

As entradas de cada RNA são um conjunto de valores de fluxo correspondentes a posições do comprimento de onda, contidos em intervalos definidos originalmente por Módolo et al. (2016). Os intervalos se referem às regiões em que os elementos H, Si, S e He se manifestam e são as mesmas regiões analisadas por especialistas humanos. Para extrair as entradas, primeiramente, cada espectro passa por uma interpolação a cada 8 pontos começando em 4000Å até 7000Å, resultando em 375 pontos. Desses 375 pontos uma quantidade de pontos é selecionada para cada

RNA que apresentam conjunto de entradas diversos. As entradas são as seguintes:

- 5000 a 6500Å: para classificar em tipo Ia ou não;
- 5500 a 7000Å: para classificar em tipo Ib ou não;
- 5500 a 6500Å: para classificar em tipo Ic ou não;
- 4000 a 5000Å e 6000 a 7000Å: para classificar em tipo II ou não.

Cada espectro origina quatro conjuntos de entrada, um para cada RNA. A Figura 4 mostra uma representação gráfica desses conjuntos para um espectro da SN2002bo.

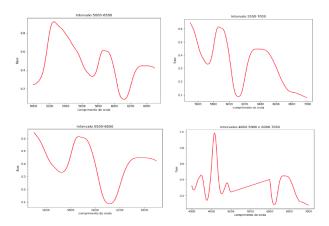


Figura 4: Representação gráfica dos Intervalos escolhidos como entradas das 4 RNAs.

5.2 Parâmetros de treinamento

Alguns parâmetros numéricos são necessários para executar o treinamento de uma rede neural. A taxa de aprendizado e a taxa de momentum são responsáveis pela estabilidade do algoritmo retropropagação. O bias é um fator que serve para diminuir ou aumentar a saída do somador de cada neurônio. A função de ativação é a função que define a saída de cada neurônio. Além disso, é preciso definir uma taxa de erro que será tolerado pelo algoritmo e os pesos iniciais. A quantidade máxima de épocas que devem ser executadas foi estabelecida como 250000. Todos esses parâmetros, exceto pela quantidade máxima de épocas, foram definidos por Módolo et al. (2016) na primeira versão da CIntIa, classificador que este trabalho se propõe a aperfeiçoar, e são eles:

- Taxa de aprendizado = 0.5;
- Taxa de Momentum = 0.3;
- Bias = 1;
- Erro tolerado = 0.001;
- Função de ativação = sigmóide;
- Pesos iniciais: escolhidos aleatoriamente.

Topologias testadas

Executou-se 7 experimentos para cada RNA, como mostrado na Tabela 1, a fim de escolher a topologia mais adequada para cada uma delas. O máximo de camadas ocultas permitidas pelo sistema são duas.

Tabela 1: Topologias testadas para escolha da mais adequada para cada RNA da CINTIA 2.

Topologia	Neur. camada 1	Neur. camada 2		
1	5	0		
2	10	0		
3	25	0		
4	40	0		
5	10	10		
6	20	4		
7	40	8		

Neur. camada 1: número de neurônios na primeira camada.

Neur. camada 2: número de neurônios na segunda camada.

Resultados e discussão

Os resultados alcançados neste trabalho redundam no aperfeiçoamento da primeira versão da CIntIa chamada agora de CINTIA 2. As topologias das quatro RNAs que compõe o sistema foram ajustadas e usando os resultados individuais propomos uma arquitetura hierárquica para classificar as SNs sem incorrer em ambiguidades. A partir da arquitetura, desenvolvemos uma ferramenta usando as linguagem de programação Python e C++, que em breve estará disponível para os interessados. Nesta seção, tratamos dos resultados dos testes individuais para cada RNA binária, da arquitetura da CINTIA 2 e da ferramenta.

6.1 Resultados das RNAs binárias

Nesta seção, agrupamos os resultados das 4 RNAs binárias. A Tabela 2 exibe a quantidade de espectros usada nos testes (40% dos espectros), por tipo. A quantidade é ligeiramente diferente porque os treinamentos e testes de cada RNA foram feitos independentemente dos outros. Todos os espectros das RNAs Ib, Ic e II não tem limitação de fase espectral, enquanto os espectros do tipo Ia usados na RNA Ia variam de -10.9 dias a +10.9 dias.

Tabela 2: Quantidade de espectros usados nos testes das 4 RNAs.

		•		
	Espectros Ia	Espectros Ib	Espectros Ic	Espectros II
RNA Ia RNA Ib RNA Ic RNA II	1405 2588 2582 2586	180 178 179 180	212 212 212 212	592 590 591 594

Foram realizados 2 testes para cada RNA, consideramos aqui os melhores resultados. Apresentamos na Tabela 3 as métricas de avaliação mais usadas para classificadores binários, seguindo o relato de Sokolova and Lapalme (2009). Para avaliação dos resultados, também é usado o Índice Kappa que mede a concordância entre duas ou mais classificação, segundo Landis and Koch (1977).

6.2 Arquitetura da CINTIA 2

Os resultados dos testes realizados em cada uma das RNAs binárias, individualmente, levaram ao desen-

Tabela 3: Métricas de avaliação da classificação das 4 RNAs binárias.

	RNA Ia	RNA Ib	RNA Ic	RNA II
Acurácia	0.92	0.97	0.95	0.96
Precisão	0.99	0.87	0.85	0.97
Recall	0.87	0.58	0.21	0.77
F1-Score	0.93	0.70	0.35	0.86
Especificidade	0.99	0.99	0.99	0.99
AUC	0.93	0.79	0.61	0.88
Índice Kappa	0.84	0.69	0.33	0.84

volvimento de uma arquitetura para a CINTIA 2. Ressaltamos que uma maneira de integrar as RNAs não havia sido proposta na primeira versão da CIntIa. A importância dessa arquitetura é possibilitar o desenvolvimento de um sistema classificador que não produza respostas ambíguas. No entanto, uma primeira versão da hierarquia foi proposta por do Nascimento et al. (2019), nessa versão preliminar a ordem dos módulos da hierarquia é ligeiramente diferente. A melhoria apresentada aqui garante mais qualidade na classificação, as métricas de avaliação são superiores.

Cada uma das RNAs é um módulo na visão geral do classificador, a Figura 5 mostra as topologias das RNAs em cada módulo. Enquanto a Figura 6 apresenta a visão geral da arquitetura da CINTIA 2. A hierarquia foi definida seguindo os valores de Índice Kappa, o módulo com maior índice é consultado primeiro. No caso dos módulos Ia e II, que apresentam o mesmo valor de Kappa o desempate foi feito pelos valores de F1-Score, índice que relaciona a Precisão e o Recall, duas métricas importantes para auferir a qualidade do método utilizado.

O fluxo da classificação inicia identificando se o espectro, previamente pré-processado, é do tipo Ia ou não, privilegiando a identificação do tipo considerado mais importante para os estudos atuais de Cosmologia. Em seguida, são classificados os do tipo II, único tipo que apresenta linhas espectrais de H. Na sequência os espectros devem ser classificados como tipo Ib ou tipo Ic, que é a diferenciação mais desafiadora a ser feita pelo sistema inteligente, assim como é para os especialistas humanos. Caso o espectro não seja classificado em nenhum dos módulos ele recebe o rótulo de "Tipo Não Identificado". A Tabela 4 mostra as métricas de avaliação da CINTIA 2 considerando o sistema como um classificador multi-classes, as métricas foram calculadas de acordo com Sokolova and Lapalme (2009)

A Acurácia M (média) tem um excelente resultado, pois a acurácia dos módulos individuais também é muito boa. A Taxa de Erro é baixa, mas os erros que acontecem são majoritariamente de falsos negativos, logo isso se reflete nos valores de Recall. Apesar disso, o F1-Score (principalmente o μ) confere ao classificador um bom índice.

6.3 Ferramenta CINTIA 2

A ferramenta CINTIA 2 foi desenvolvida de acordo com a arquitetura apresentada na subseção anterior. Além das 4 redes neurais binárias ela possui um módulo de pré-processamento que prepara os espectros para a classificação. As linguagens usadas no desenvolvimento foram Python e C++. Com Python

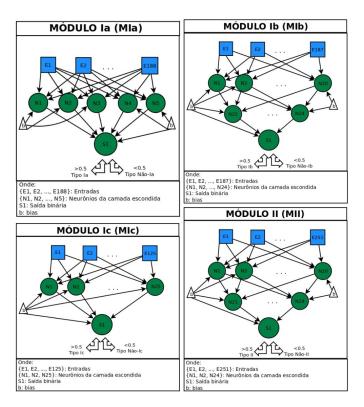


Figura 5: Módulos que compõem a CIINTIA 2. MIa classifica em tipo Ia ou não-Ia; MIb classifica em tipo Ib ou não-Ib; MIc classifica em tipo Ic ou não-Ic; MII classifica em tipo II ou não-II.

Tabela 4: Métricas de avaliação da CINTIA 2, depreendidos dos resultados dos testes dos módulos individuais.

Acurácia M	Taxa de Erro	Precisão M	Recall M	F1-Score M	Precisão μ	Recall μ	F1–Score μ
0.95	0.05	0.92	0.61	0.73	0.98	0.77	0.86

desenvolvemos o pré-processamento e um módulo simples para visualização de dados. As redes neurais foram desenvolvidas em C++ visando a manutenção de componentes de software produzidos na primeira versão da CIntIa. A tela que inicia a execução do classificador é um prompt de comando. A CINTIA 2 pede que seja informado o caminho de um arquivo (.csv) com valores dos redshifts das SNs cujos espectros devem ser classificados. Em seguida, pede-se que seja informado o caminho de uma pasta onde estão arma-

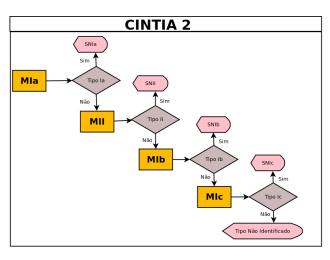


Figura 6: Arquitetura da CINTIA 2.

zenados os espectros que devem ser classificados.

O módulo de visualização de dados apresenta duas funções simples, mas que podem ajudar na confirmação do tipo classificado pela CINTIA 2. Todos os espectros são impressos na tela e em seguida, caso haja 10 espectros, ou mais, na pasta para serem classificados, são impressos 4 gráficos de dispersão. Cada gráfico de dispersão, um por rede neural, apresenta a distribuição espacial das duas componentes principais, calculadas pelo método de Isomap, possibilitando assim a comparação entre os espectros considerando apenas as entradas de cada rede neural. A Figura 7 mostra os gráficos exibidos pela CINTIA 2.

A classificação é apresentada na tela, também em prompt de comando, e salva em um arquivo (.csv). A Figura 8 mostra a tela com as classificações feitas pela CINTIA 2.

6.4 Comparação com outros classificadores

Uma comparação entre a CINTIA 2 e os outros classificadores consultados na literatura é importante para avaliar em que aspectos ela representa uma melhoria e depreender o quão vantajosa é a sua utilização. A Tabela 5 apresenta os critérios escolhidos para efetuar essa confrontação.

Considerando a quantidade de dados e a abrangência das fases espectrais, a CINTIA 2 é superior a todos os outros classificadores, ou seja, ela compreende uma diversidade maior de espectros. A CINTIA 2

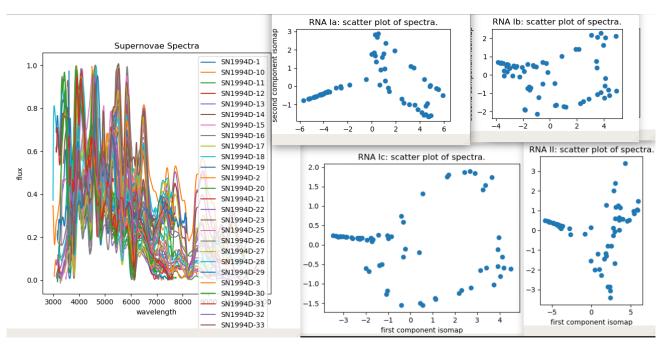


Figura 7: Visualização de dados exibida pela CINTIA 2. O primeiro quadro (esquerda) exibe todos os espectros da pasta. Os outros 4 quadros são gráficos de dispersão das entradas de cada uma das RNAs, depois que passam pela redução de dimensionalidade usando o método Isomap.

Tabela 5: Métricas de avaliação que comparam a CINTIA 2 com a CIntIa e os outros classificadores da Seção 3.

	-	•	-	-				•	_
Classificador		SNID	GELATO	CIntIa	DRACULA	Quantitative Classification	SUZAN	CINTIA 2	
	Espectros Ia	879	1009	559	486	146	3082	6644	
	Espectros Ib	NC	129	33	NC	12	217	461	
	Espectros Ic	NC	355	44	NC	19	282	538	
	Espectros Ib/c	322	83	NC	NC	4	NC	NC	
	Espectros II	353	1132	13	NC	NC	116	1513	
Me	nor Fase Espectral	-15	-15	-3	-3	-5	-15	-1435	
Ma	aior Fase espectral	+70	+600	+7	+3	+5	+2959	+8303	
	Acurácia	NC	NC	0.98	NC	0.95	0.73	0.95	
	Precisão	NC	NC	0.97	NC	NC	0.93	0.97	
	Recall	NC	NC	0.94	NC	NC	0.72	0.76	
	F1-Score	NC	NC	0.95	NC	NC	0.83	0.85	

Os valores NC indicam que a informação não consta no trabalho. Para as duas versão da CIntIa, as métricas Acurácia, Precisão, Recall e F1-Score se referem aos valores μ .

transcende todos os classificadores consultados, que apresentam bons resultados apenas para espectros obtidos durante o brilho máximo da SN. Considerando que a expectativa é tornar o sistema um classificador operacional em conjunto com telescópios, essa é uma característica importante. A sua Acurácia Média e a Precisão μ são equiparáveis aos outros sistemas que também apresentam essa informação. Em relação à primeira versão da CIntIa, as métricas de avaliação da CINTIA 2, consideradas na Tabela 4, são inferiores. No entanto, a quantidade de dados melhora a análise proposta, evidenciando a capacidade da hierarquia de RNAs binárias em aprender padrões dos tipos Ia e II e a dificuldade em diferenciar espectros de SNs Ib e Ic, as quais têm a mesma origem física.

Conclusão

O trabalho descrito neste artigo, apresenta a CINTIA 2, que é o aprimoramento do Classificador Inteligente de SNs do tipo Ia, proposto originalmente por Módolo et al. (2016). A primeira CIntIa foi treinada com uma quantidade limitada de dados, provenientes de apenas uma base de espectros dentre as disponíveis gratuitamente. A atualização desse sistema é uma necessidade vigente, pois cada vez mais cresce a quantidade de telescópios varrendo os céus em busca de SNs. Nesse contexto, a CINTIA 2 apresenta melhorias em relação ao classificador original. Aumentou-se a quantidade e variedade de dados, a amplitude das fases espectrais foi alargada e os resultados, consequentemente, são bons. A CINTIA 2 faz distinção, com excelência, entre espectros de SNs do tipo Ia e os outros tipos. As SNs do tipo II também são adequadamente reconhecidas pelo classificador. A classificação do tipo Ib tem bons resultados, mas não excelentes,

enquanto a classificação do tipo Ic é apenas regular. Uma comparação com outros sistemas relatados na literatura atestam a qualidade da CINTIA 2.

Outro progresso, alcancado neste trabalho, é o desenvolvimento de uma arquitetura que conecta as RNAs binárias individuais (módulos) em um sistema hierárquico completo, o que ainda não havia sido proposto para a primeira versão e é um avanço na garantia de que a classificação não será ambígua. Todos os espectros recebem apenas um rótulo identificando o tipo da SN a qual pertencem, ao fim da execução do sistema. A CINTIA 2 foi desenvolvida usando as linguagens de programação Python e C++, garantido livre acesso (a ferramenta estará disponível em breve) e portabilidade do software.

Entendemos que as contribuições deste trabalho são: a inclusão de uma grande quantidade de dados no treinamento das RNAs, o que permite uma análise robusta do desempenho; a proposta de arquitetura de hierarquia das redes neurais binárias que ainda não existia na CIntIa, assim a classificação não apresenta ambiguidade; e o desenvolvimento prático da ferramenta que está pronta para uso em conjunto com quaisquer telescópios e espectrógrafos.

Futuramente, espera-se definir as topologias e outros parâmetros de cada módulo da CINTIA 2 automaticamente, assim é possível melhorar a qualidade da classificação, já que mais testes serão feitos em um tempo menor, tarefa árdua para ser realizada manualmente.

```
Generating files...
Classifying...
Running module II...
Running module Ia...
Running module Ib...
Running module Ic...
 -----CLASSIFICATION-----
Pattern: 1
Supernova: SN1994D
Spectrum: SN1994D-1
Classification: Tipo Ia
Pattern: 2
Supernova: SN1994D
Spectrum: SN1994D-10
lassification: Tipo Ia
Pattern: 3
Supernova: SN1994D
Spectrum: SN1994D-11
Classification: Tipo Ia
Pattern: 4
Supernova: SN1994D
Spectrum: SN1994D-12
.
Classification: Tipo Ia
Pattern: 5
Supernova: SN1994D
Spectrum: SN1994D-13
Classification: Tipo Ia
```

Figura 8: Exibição dos resultados de classificação da CINTIA 2 em prompt de comando.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento

Referências

Anad, R., Mehrotra, K., Mohan, C. and Ranka, S. (1995). Efficient classification for multiclass problems using modular neural networks, IEEE Transactions on Neural Networks 6(1): 117-124. https: //doi.org/10.1109/72.363444.

Arantes Filho, L. R. (2018). Classificação inteligente de supernovas utilizando sistemas de regras nebulosas, Dissertação (mestrado em computação aplicada), Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos.

Blondin, S., Mandel, K. S. and Kirshner, R. P. (2011). Do spectra improve distance measurements of type ia supernovae?, Astronomy & Astrophysics **526**(A81). https://doi.org/10.1051/0004-6361/201015792.

Blondin, S., Matheson, T., Kirshner, R. P., Mandel, K. S., Berlind, P., Calkins, M., Challis, P., Garnavich, P. M., Jha, S. W., Modjaz, M., Riess, A. G. and Schmidt, B. P. (2012). The spectroscopic diversity of type ia supernovae, The Astronomical Journal 143(5): 126. https://doi.org/10.1088/ 0004-6256/143/5/126.

Blondin, S. and Tonry, J. L. (2007). Determining the type, redshift, and age of a supernova spectrum, The Astrophysical Journal 666(2). https://doi.org/ 10.1086/520494.

Britannica, E. (2018). Redshift. Available at https: //www.britannica.com/science/redshift (Acessed 8 January 2019).

Carroll, B. W. and Ostlie, D. A. (2007). An Introduction to Modern Astrophysics, Pearson Addison-Wesley, San Francisco. 1358 p.

Damineli, A. and Steiner, J. (2010). Fascínio do Universo, Odysseus Editora, São Paulo. 120 p.

do Nascimento, F. J. B., Arantes Filho, L. R. and Guimarães, L. N. F. (2019). Intelligent classification of supernovae using artificial neural networks, Inteligencia Artificial 22(63): 39-60. https://doi.org/ 10.4114/intartif.vol22iss63pp39-60.

Ferneda, E. (2006). Redes neurais e sua aplicação em sistemas de recuperação de informação, Ci. Inf. Brasília 35(1): 25-30. http://dx.doi.org/10.1590/ S0100-19652006000100003.

Gillochon, J., Parrent, J., Kelley, L. Z. and Margutti, R. (2017). An open catalog for supernova data, The Astrophysical Journal 835(1): 64. https://doi.org/10. 3847/1538-4357/835/1/64.

Harutyunyan, A. (2008). Automatic Objective Classification of Supernovae, PhD thesis, Universitá degli Studi de Padova.

- Haykin, S. (2001). Redes neurais: princípios e práticas, Bookman, Porto Alegre.
- Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data, Biometrics 33(1): 159-174. https://doi.org/10.2307/ 2529310.
- Luger, G. F. (2013). Inteligência Artificial, Pearson Education do Brasil, São Paulo.
- Matheson, T., Kirshner, R. P., Challis, P., Jha, S., Garnavich, P. M., Berlind, P., Calkins, M. L., Blondin, S., Balog, Z., Bragg, A. E., Caldwell, N., Concannon, K. D., Falco, E. E., Graves, G. J. M., Huchra, J. P., Kuraszkiewicz, J., Mader, J. A., Mahdavi, A., Phelps, M., Rines, K., Song, I. and Wilkes, B. J. (2008). Optical spectroscopy of type ia supernovae, The Astronomical Journal 135(4). https: //doi.org/10.1088/0004-6256/135/4/1598.
- Módolo, M., Guimarães, L. N. F. and Rosa, R. R. Supernovae automatic classification method by modeling human analysis using artificial neural networks, 2016 International Conference on Nonlinear Science and Complexity (NSC).
- Mencia, E. and Furnkranz, J. (2008). Pairwise learning of multilabel classifications with perceptrons, 2008 IEEE International Joint Conference on Neural Networks. https://doi.org/10.1109/IJCNN.2008. 4634206.
- Modjaz, M., Blondin, S., Kirshner, R. P., Matheson, T., Berlind, P., Bianco, F. B., Calkins, M. L., Challis, P., Garnavich, P., Hicken, M., Jha, S., Liu, Y. Q. and Marion, G. H. (2014). Optical spectra of 73 stripped-envelope core-collapse supernovae, The Astronomical Journal 147(5). https://doi.org/10. 1088/0004-6256/147/5/99.
- Oong, T. and Isa, N. (2012). One-against-all ensemble for multiclass pattern classification, Applied Soft Computing 12(4): 1303-1308. https://doi.org/10. 1016/j.asoc.2011.12.004.
- Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J. and Project, T. S. C. (1999). Measurements of omega and lambda from 42 high-redshift supernovae, The Astrophysical Journal **517**(2): 565. https://doi.org/10.1086/307221.
- Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilli-land, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B. and Tonry, J. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant, The Astronomical Journal 116(3): 1009. https://doi.org/10.1086/300499.

- Sasdelli, M., Ishida, E. E. O., Vilalta, R., Aguena, M., Busti, V. C., Camacho, H., Trindade, A. M. M., Gieseke, F., de Souza, R. S., Fantaye, Y. T. and Mazzali, P. A. (2016). Exploring the spectroscopic diversity of type ia supernovae with dracula: a machine learning approach, Monthly Notices of the Royal Astronomical Society 461(2): 2044–2059. https://doi.org/10.1093/mnras/stw1228.
- Savitzky, A. and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures, Analytical Chemistry 36(8): 1627-1639. https://doi.org/10.1021/ac60214a047.
- Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks, Information Processing and Management **45**: 427-437. https://doi.org/10.3847/1538-4357/ 835/1/64.
- Sun, F. and Gal-Yam, A. (2017). Quantitative classification of type i supernovae using spectroscopic features at maximum brightness, ArXiv e-prints. Available at https://arxiv.org/abs/1707.02543 (Acessed 18 April 2018).
- Turatto, M. (2003). Classification of Supernovae, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 21-
- Zhu, Y., Wang, L., Xiangyan, Y., Gu, B., Li, X., Yang, S., Gong, X., Du, F., Qi, Y. and Xu, L. (2014). Kunlun dark universe survey telescope, Vol. 9145, p. 9145. https://doi.org/10.1117/12.2055768.