

Diákok teljesítményének osztályozása életviteli mutatók alapján

Marsó Máté
NKDG9I
Óbudai Egyetem
Alba Regia Műszaki Kar
Székesfehérvár, Magyarország
marso.mate@stud.uni-obuda.hu

Pásztóhy László Ádám
H3A8UV
Óbudai Egyetem
Alba Regia Műszaki Kar
Székesfehérvár, Magyarország
pasztohylaszlo@stud.uni-obuda.hu

Ez a dokumentum diákok tanulási- és munkabéli teljesítményének, koncentrációjának alakulását vizsgálja az életükben történő különböző események, és mutatók alapján. Az ezen eseményeket és mutatókat tartalmazó adathalmazzal feltanításra került egy neurális háló, mely ezáltal képes megjósolni egy tanuló jövőbeli teljesítményét a meghatározott események alapján. Kutatásunk célja, hogy elemezni tudjuk, hogy miként változik a diákok teljesítménye az életbeli mutatóik alapján, valamint hogy a neurális háló segítségével jóslatokat tudjunk tenni ezen teljesítmény alakulására.

tanulás, neurális háló, életvitel, diákok

I. BEVEZETÉS

A projekt célja egy teljesítmény prediktáló osztályozó neurális háló felépítése, mely a kapott adatok elemzésével megadja az aznapi teljesítményt. Ezzel szeretnénk egy olyan szoftvert létrehozni, amivel egy életviteli mutatókból álló adathalmaz alapján osztályozni tudjuk a diákok teljesítményét. A kutatásunkban keressük az optimális állapotot a legjobb teljesítmény eléréséhez. A projekt prediktív mivoltja azt takarja, hogy a kutatás végeztével birtokunkban lesz egy olyan program, mely képes a fentebb említett mutatók megadása után előrejelezni a lehetséges teljesítményt. A motivációnk a diákok szokásainak teljesítményükre gyakorolt hatásának feltérképezése.

II. A TÉMA BEMUTATÁSA

A. Big Data

Egy olyan kifejezés, amelyet egy hatalmas mennyiségű strukturált és/vagy strukturálatlan adatra használnak, amelyből ismereteket nyernek ki. Ennek a mennyiségnek olyan óriási méretei vannak, hogy nagyon nehéz feldolgozni őket a hagyományos adatbázis- és szoftvertechnikák segítségével. [1]

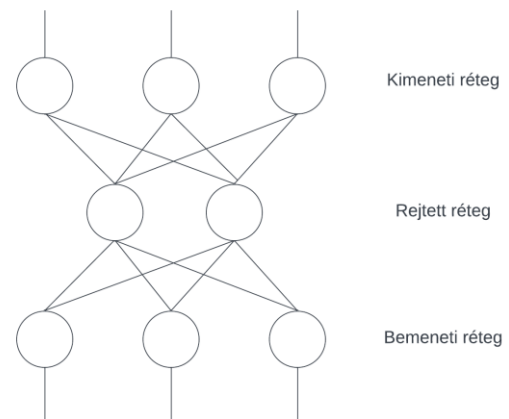
B. Neurális hálók

A neurális hálók az emberi agy matematikai szimulációi. Az emberi agyban megközelítőleg 10^{11} neuronból áll, melyek kémiai reakciók során kommunikálnak egymással (neurotranszmitterek). Bizonyos bemenetek hatására az adott neuron – ha azok átlélik annak ingerküszöbét – kimenő jelet küld a körülötte álló neuronoknak, így létrehozva egy kommunikációs hálót.

Ennek gépi szimulációja, a mesterséges neuron n bemenetet hatására produkál egy kimenetet (vagy kimenetek sorozatát). Ezek a bemenetek érkezhettek a hálón kívülről, vagy a hálóban szereplő többi, szomszédos neurontól. Egy neuronnak lehet aktiválási függvényt meghatározni, mely meghatározza, hogy milyen erősségű bemeneti jel szükséges

az adott neuron aktiválásához. Ezzel kiküszöbölhető a neuronok túlérzékenysége.

Ezek a neuronok rétegekbe vannak szervezve, és ezáltal kapcsolódnak egymáshoz. Általában létezik egy bemeneti réteg, valamint egy kimeneti réteg. Ezek között találhatóak meg a rejtett rétegek. Az összes réteg együttesen alkotja a neurális hálót, melynek több változata is van: létezik generatív háló, mely létrehoz adatokat (szöveget, képeket, stb...), valamint létezik osztályozó háló, mely a kapott adathalmazt dolgozza fel, és a megtanult adatok alapján sorolja osztályokba az újonnan beérkező adatokat. [2]



1. ábra: A legegyszerűbb, háromrétegű neurális háló felépítése [2]

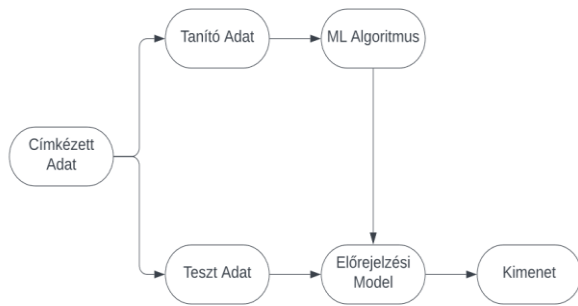
Ebben a kutatásban egy osztályozó hálót fogunk használni, melynek tanítási folyamatát, valamint az eredmények kiértékelését tartalmazza ezen dokumentum.

C. Felügyelt tanulás

A gépi tanulási algoritmusok minden adatpontot az adatkészletben (Dataset) tulajdonságok gyűjteményeként kezelnek. [3]

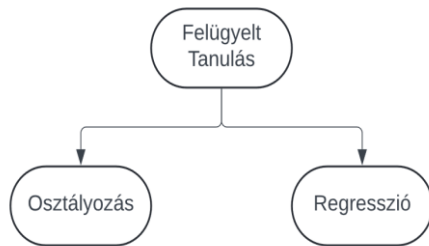
Ha az adatpontok címkézettek, akkor ezt a tanulási típust felügyelt tanulásnak nevezzük. A felügyelt tanulás a modellek tanítását jelenti címkézett adatokon, majd a tesztelésüket címkézetlen adatokon. [3]

A felügyelt tanulás alapvető architektúrája a következő lépéseket foglalja magában: adathalmaz gyűjtése, az adathalmaz felosztása tesztelési és tanítási adatokra, majd az adat előfeldolgozása. A kinyert tulajdonságokat beillesztik egy algoritmusba, majd a modellt megtanítják azoknak a tulajdonságoknak a megtanulására. [3]



2. ábra: A felügyelt tanulás alapvető architektúrája [3]

Az osztályozás és a regresszió a felügyelt tanulás két nagy típusa. A következőkben az osztályozással fogunk foglalkozni.



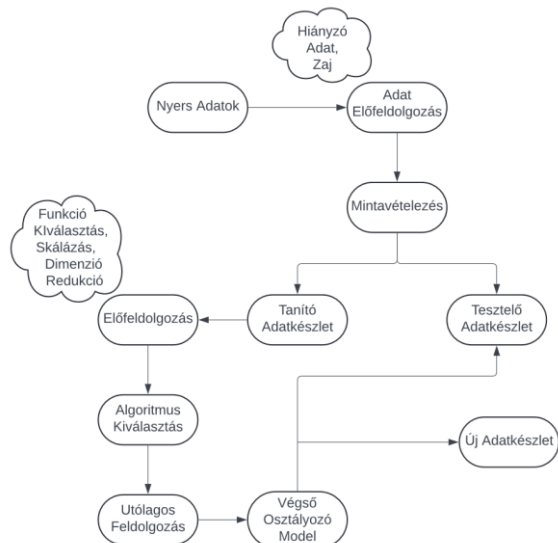
3. ábra: Felügyelt tanulás típusai [3]

D. Osztályozási modell

A felügyelt osztályozás az úgynevezett Intelligens Rendszerek által leggyakrabban végzett feladatok egyike. [4]

A klasszifikáció során egy modell ismeretlen értékeket (kimenetek halmaza) jósol egy ismert értékek halmaza (bemenetek halmaza) alapján. Ha a kimenet kategorikus formában van, akkor a problémát klasszifikációként említjük. Általában a klasszifikáció során az adathalmaz példáit meghatározott osztályokba sorolják. [3]

Egy klasszifikációs modell egy leképzési függvényt használ, amelyet a modell a tanítóadatokból következtet, hogy megjósolja a tesztadatok osztálycímekjét. Végül egy jellemző társul az adathalmazhoz, amely segít egy pontos előre jelző modell létrehozásában. [3]



4. ábra: A felügyelt tanulásban történő osztályozás munkafolyamata [3]

III. PREPARE YOUR PAPER BEFORE STYLING

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: “Wb/m²” or “webers per square meter”, not “webers/m²”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
- Use a zero before decimal points: “0.25”, not “.25”. Use “cm³”, not “cc”. (*bullet list*)

C. Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

D. Some Common Mistakes

- The word “data” is plural, not singular.
- The subscript for the permeability of vacuum μ_0 , and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
- Do not use the word “essentially” to mean “approximately” or “effectively”.
- In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
- Do not confuse “imply” and “infer”.
- The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the “et” in the Latin abbreviation “et al.”.
- The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

IV. USING THE TEMPLATE

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

A. Authors and Affiliations

The template is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

1) *For papers with more than six authors:* Add author names horizontally, moving to a third row if needed for more than 8 authors.

2) *For papers with less than six authors:* To change the default, adjust the template as follows.

a) *Selection:* Highlight all author and affiliation lines.

b) *Change number of columns:* Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

c) *Deletion:* Delete the author and affiliation lines for the extra authors.

B. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

C. Figures and Tables

a) *Positioning Figures and Tables:* Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

^a Sample of a Table footnote. (*Table footnote*)

Fig. 1. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

ACKNOWLEDGMENT (*Heading 5*)

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

TABLE I. TABLE TYPE STYLES

Table Head	Table Column Head		
	Table column subhead	Subhead	Subhead
copy	More table copy ^a		

V. HIVATKOZÁSOK

- [1] C. Shavari, . S. Vijender K. és J. Madhuri S., The Basics of Big Data and Security Concerns, 2017.
- [2] S. Mengel, W. Lively, Using a Neural Network to Predict Student Responses, 1992.
- [3] S. Dridi, “SUPERVISED LEARNING - A SYSTEMATIC LITERATURE REVIEW,” 2021.
- [4] K. S. B., Z. I. D. és P. P. E., „Machine learning: a review of classification,” 2007.

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

- [1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (*references*)
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [7] M. Young, *The Technical Writer’s Handbook*. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.