

Using General Large Language Models to Classify Mathematical Documents

Patrick D.F. Ion *and* Stephen M. Watt

17th Conference on Intelligent Computer Mathematics
Montreal, August 5–9 2024

Outline

1 Introduction

2 Example

3 Methods

4 Results

5 Discussion

6 Conclusions & Future

About Large Language Models

- What is all this AI hype?
- Can chatbots be used for classification without further ado: even MSC?
- What is it they do? How does one prompt?
- Is a resulting MSC suggestion reliable, and how or why?
- How much input is needed? title, abstract, paper, references?
- Does the actual LLM used matter?
- Can it all be automated? And used by ZMO and MSN?
- Does this say anything about math and truth?
- Therefore start by hand.

Mathematics Subject Classification

00 General and overarching topics; collections	45 Integral equations
01 History and biography	46 Functional analysis
03 Mathematical logic and foundations	47 Operator theory
05 Combinatorics	49 Calculus of variations and optimal control; optimization
06 Order, lattices, ordered algebraic structures	51 Geometry
08 General algebraic systems	52 Convex and discrete geometry
11 Number theory	53 Differential geometry
12 Field theory and polynomials	54 General topology
13 Commutative algebra	55 Algebraic topology
14 Algebraic geometry	57 Manifolds and cell complexes
15 Linear and multilinear algebra; matrix theory	58 Global analysis, analysis on manifolds
16 Associative rings and algebras	60 Probability theory and stochastic processes
17 Nonassociative rings and algebras	62 Statistics
18 Category theory; homological algebra	65 Numerical analysis
19 K -theory	68 Computer science
20 Group theory and generalizations	70 Mechanics of particles and systems
22 Topological groups, Lie groups	74 Mechanics of deformable solids
26 Real functions	76 Fluid mechanics
28 Measure and integration	78 Optics, electromagnetic theory
30 Functions of a complex variable	80 Classical thermodynamics, heat transfer
31 Potential theory	81 Quantum theory
32 Several complex variables and analytic spaces	82 Statistical mechanics, structure of matter
33 Special functions	83 Relativity and gravitational theory
34 Ordinary differential equations	85 Astronomy and astrophysics
35 Partial differential equations	86 Geophysics
37 Dynamical systems and ergodic theory	90 Operations research, mathematical programming
39 Difference and functional equations	91 Game theory, economics, social and behavioral sciences
40 Sequences, series, summability	92 Biology and other natural sciences
41 Approximations and expansions	93 Systems theory; control
42 Harmonic analysis on Euclidean spaces	94 Information and communication, circuits
43 Abstract harmonic analysis	97 Mathematics education
44 Integral transforms, operational calculus	

Mathematics Subject Classification

- 00 General and overarching topics; collections
- 01 History and biography
- 03 Mathematical logic and foundations
- 05 Combinatorics
- 06 Order
- 08 General
- 11 Number theory
- 12 Field theory
- 13 Commutative algebra
- 14 Algebraic geometry
- 15 Linear and multilinear algebra; matrix theory
- 16 Associative rings and algebras
- 17 Nonassociative algebras
- 18 Category theory and homological algebra
- 19 K -theory
- 20 Group theory and generalizations
- 22 Topological groups
- 26 Real functions
- 28 Measure theory, integration, and probability theory
- 30 Function theory
- 31 Potential theory
- 32 Several complex variables and analytic spaces
- 33 Special functions
- 34 Ordinary differential equations
- 35 Partial differential equations
- 37 Dynamical systems and ergodic theory
- 39 Difference and functional equations
- 40 Sequences, series, summability
- 41 Approximations and expansions
- 42 Harmonic analysis on Euclidean spaces
- 43 Abstract harmonic analysis
- 44 Integral transforms, operational calculus
- 45 Integral equations
- 46 Functional analysis
- 47 Operator theory
- 49 Calculus of variations and optimal control; optimization

51 Geometry

52 Convex and discrete geometry

53 Differential geometry

54 General topology

55 Algebraic topology

57 Manifolds and cell complexes

58 Global analysis, analysis on manifolds

- 81 Quantum theory
- 82 Statistical mechanics, structure of matter
- 83 Relativity and gravitational theory
- 85 Astronomy and astrophysics
- 86 Geophysics
- 90 Operations research, mathematical programming
- 91 Game theory, economics, social and behavioral sciences
- 92 Biology and other natural sciences
- 93 Systems theory; control
- 94 Information and communication, circuits
- 97 Mathematics education

Mathematics Subject Classification

New MSC 2020 (19 June 2023 version) contains:

- 63 2-digit classifications, e.g. [53](#) (Geometry)
- 529 3-digit classifications, e.g. [53D](#) (Symplectic & Contact G.)
- 6,022 5-digit classifications, e.g. [53D25](#) (Geodesic Flows in S&C G.)
- Full text of a 2-digit classification suggests related, e.g. [53-XX](#) *Differential geometry For differential topology, see 57Rxx; for foundational questions of differentiable manifolds, see 58Axx*

More of a DAG than a tree.

Recent zbMATH articles in MSC 53; 145,445 items

- Descriptions of classes were changed in 2020 to be more useful when searching online or via database interfaces.

Outline

1 Introduction

2 Example

3 Methods

4 Results

5 Discussion

6 Conclusions & Future

ChatGPT 3.5

- Generative Pre-Trained Transformer with NLP front-end
- Offered free by OpenAI

Example

ChatGPT 3.5 ▾



You

Call this conversation "Math Classification by LLM".



ChatGPT

Sure, I've noted the conversation title as "Math Classification by LLM." How can I assist you further with math classification or any other related topics?

Example



You

Remember the following abstract for future use and call it "AdamsAbstract":
The Adams conjecture states that the local theta correspondence sends a local Arthur packet to another local Arthur packet. Mœglin confirmed the conjecture when lifting to groups of sufficiently high rank and also showed that it fails in low rank. Recently, Bakić and Hanzer described when the Adams conjecture holds in low rank for a representation in a fixed local Arthur packet. However, a representation may lie in many local Arthur packets and each gives a minimal rank for which the Adams conjecture holds. In this paper, we study the interplay of intersections of local Arthur packets with the Adams conjecture.



ChatGPT

Got it! The abstract has been saved with the title "AdamsAbstract" for future reference. If you need to refer to it or discuss its content further, just let me know!

Example



You

Classify "AdamsAbstract" according to the MSC 2020 classification.



ChatGPT

The Mathematical Subject Classification (MSC) is a system used to categorize mathematical research papers. Let's classify "AdamsAbstract" according to the MSC 2020 classification. Based on the content provided, it seems to fall under the category of representation theory. Here's a classification based on the provided abstract:

Example

Primary classification:

- 22E50: Representations of Lie and linear algebraic groups over local fields

Secondary classifications:

- 11F27: Theta series; Weil representation; theta correspondences
- 20G25: Linear algebraic groups over local fields and their integers
- 11F70: Representation-theoretic methods; automorphic representations over local and global fields

Example

This classification is based on the topics mentioned in the abstract, such as the local theta correspondence, representation theory of Lie groups over local fields, and the study of Arthur packets. If you need further classification or clarification, feel free to ask!

Example



You

What additional MSC 2020 areas are relevant?



ChatGPT

In addition to the primary and secondary classifications provided earlier, here are some additional MSC 2020 areas that may be relevant based on the content of the abstract:

Example

- 22E45: Representations of Lie and linear algebraic groups over real fields: analytic methods
- 22E55: Representations of Lie and linear algebraic groups over global fields and adèle rings
- 22E65: Infinite-dimensional Lie groups and their representations
- 11F41: Automorphic forms on $GL(2)$; Hilbert and Hilbert-Siegel modular groups and their modular and automorphic forms; Hilbert modular surfaces
- 11F70: Representation-theoretic methods; automorphic representations over local and global fields

Example

These additional areas cover topics related to representations of Lie groups over real and global fields, as well as infinite-dimensional Lie groups, automorphic forms, and modular forms, which might also be relevant depending on the specific content of the paper. If you need further assistance or clarification, feel free to ask!

Outline

1 Introduction

2 Example

3 Methods

4 Results

5 Discussion

6 Conclusions & Future

Sample

- Took the most recent article from the arXiv for each MSC 2020 top-level classification, except for the new 97 (Math Ed).
- Took most recent to ensure articles not part of training set.
- Used arXiv search under “all fields”, not just “mathematics”
- Withdrawn papers not considered.
- Some articles were the most recent in more than one classification.
- Resulting sample had 56 articles.

- For each article, the following query was processed:

```
Call the following text "<arXivId>-Title":  
<title from paper>
```

```
Call the following text "<arXivId>-Abstract":  
<abstract from paper>
```

```
Given the title "<arXivId>-Title" and abstract  
"<arXivId>-Abstract" classify the text according to the  
MSC 2020 classification.
```

- The output for each was similar to our earlier example.

Analysis

- In each case, we determined whether the LLM classification matched the arXiv-ascribed classification of the article.
- If **all** of the primary classifications offered by the LLM chat were among the primary arXiv classifications, it was taken to be **matching**. We allow to miss arXiv primary classifications.
- If **any** of the primary classifications offered by the LLM chat were **not** among the primary arXiv classification, it was taken to be **differing**.

Outline

1 Introduction

2 Example

3 Methods

4 Results

5 Discussion

6 Conclusions & Future

Matching and Differing Classifications

- 34 of the 56 papers had matching classifications.
- 22 of the 56 papers had differing classifications.

Matching Classifications

MSC Section	arXiv ID	arXiv MSC	LLM primary	LLM secondary	# arXiv MSC			
					# LLM primary	# MSC	# primary missed	# secondary extra
00	2403.16849	(00; 97; 97A99; 97A40)	00A05	97U50	2	1	-	-
01	2311.16649	(01)	01Axx	01A55, 01A60; 30Exx, 30Fxx, 30Gxx	1	1	-	1
03	2403.07616	(03)	03Cxx	03C45, 03C13, 03C98	1	1	-	-
05	2401.13436	(05; 11)	11-xx	11Pxx, 11Dxx, 11Nxx	2	1	1	-
08	2301.09966	(68; 40; 08)	68Q42	68Q45	3	1	2	-
11	2403.05453	(11; 14)	14G05, 11G25		2	2	-	-
12	9807008	(57R70; 58A5; 12; 58C35; 40 58F19; 58Q15)	58J52	53C05, 58J40, 58J50, 58J52	3	1	2	1

Matching Classifications

MSC Section	arXiv ID	arXiv MSC	LLM primary	LLM secondary				
					# arXiv MSC	# LLM primary	# MSC missed	# extra
15	2403.19691	(15)	15a42, 15A60, 15A90		1	1	-	-
16	2401.02545	(16)	16-xx	17-xx	2	1	-	1
18	2307.01498	(18)	18D05	18-02, 18D35, 18D99	1	1	-	-
19	2112.12010	(14; 13; 11; 19)	none		4	0	4	-
30	2312.13673	(30; 31)	30C15	26C10	2	1	1	1
32	2112.13212	(32; 53)	32U05, 32U15	32W20, 46G20	2	1	1	1
35	2311.15913	(34; 35; 49; 70; 74)	49M25	49K15	5	1	4	-
37	2403.13116	(37)	37-xx	37Hxx	1	1	-	-
43	1509.03761	(42; 43; 46)	42B25	46E35	3	1	1	-
45	2102.03123	(45; 74; 76; 86)	74-xx	74Fxx	4	1	3	-
47	2312.06390	(47)	47B37	47B15	1	1	-	-
52	2204.00893	(90; 05; 52; 68)	68T10	68W10	4	1	3	-
53	2312.08173	(70F; 53)	70G65	83C10	2	1	1	1
55	2308.15974	(51; 53; 55; 57K10; 58D05; 37E30)	55R65	57S30	6	1	4	-

Matching Classifications

MSC Section	arXiv ID	arXiv MSC	LLM primary	LLM secondary	# arXiv MSC			
					#	#	# primary missed	# secondary extra
58	2307.00694	(35; 53; 58)	58J50	58J32	3	1	2	-
62	2403.18245	(62)	62P20	62H12	1	1	-	-
65	2403.07875	(65)	65M12	65M55	1	1	-	-
68	2404.00549	(68)	68T45	92C55	1	1	-	-
76	2403.18088	(65 Prim; 76; 35 Sec)	76M20	65M60	1	1	-	-
78	2403.08471	(78)	78A45	78A35	1	1	-	-
80	2309.02308	(78; 80)	78A60	78M20	2	1	1	-
82	2312.14281	(82)	82D45	82C26	1	1	-	-
83	2402.13860	(83)	83-xx	83Cxx	1	1	-	-
85	2401.02337	(85)	85-xx	85A04	1	1	-	-
92	2312.12888	(90; 92)	92B20	68T05	2	1	1	-
93	2402.13772	(93)	93B35	93B07	1	1	-	-
94	2308.14725	(20; 68; 94)	94A60	20D60, 20F50, 20P05	3	1	1	-

Differing Classifications

MSC Section	arXiv ID	arXiv MSC	LLM Primary	LLM Secondary						
					# arXiv MSC	# LLM primary MSC	# LLM primary "wrong"	LLM quality	# primary missed	# secondary extra
06	2403.05604	(06)	05C15	05C20, 05C75, 05C85, 06-xx	1	1	1	=	-	1
13	2306.17679	(13; 03)	16W10	16K20, 16H05	2	1	1	+1	2	1
17	1910.03789	(17)	85-xx	85Axx	1	1	1	+2	1	1
20	2308.15765	(20)	94A60	11T71	1	1	1	+2	1	1
22	2303.01437	(22; 76)	35Q72	74B20	2	1	1	+2	2	1
26	2312.02400	(26; 40)	68-xx	68Txx	2	1	1	+2	2	-
28	1801.04970	(28)	26A39	28A12	1	1	1	-2	-	-
33	2301.05790	(33; 33Cxx)	35S05	30E20	1	1	1	=	1	1
34	2403.06996	(34; 37; 60; 62)	68T50	68T05	4	1	1	+1	4	-
39	2312.03569	(39; 81)	34A05	34L10	2	1	1	+2	2	1

Differing Classifications

MSC Section	arXiv ID	arXiv MSC	LLM Primary	LLM Secondary	# arXiv MSC		LLM quality		# primary missed		# secondary extra	
					#	#	#	#	#	#	#	#
41	2402.09991	(41)	86-xx	94A17	1	1	1	+1	1	1		
42	2311.07436	(42)	47B47	42B20	1	1	1	=	-	-		
44	2212.04345	(33; 44; 81)	42A38	81S99	3	1	1	+2	2	-		
46	2302.05234	(46)	82B44	47B80	1	1	1	+2	1	1		
51	2402.07343	(14; 51; 81)	53D12, 58J42		3	2	2	+2	3	1		
54	2303.13253	(70; 97; 54)	01	70	3	1	1	+1	2	-		
57	2403.19481	(57)	58J10	53C23	1	1	1	+1	1	1		
60	2403.15220	(60)	62-xx, 62Fxx, 62F10	68-xx, 68Txx, 68T01	1	1	1	+2	1	1		
74	2311.17485	(74)	65F90	65N99	1	1	1	+2	1	1		
86	2401.06225	(86)	68U10	62H30, 68T45	1	1	1	+2	1	-		
90	2402.12283	(90)	65K05	65K10	1	1	1	+2	1	-		
91	2402.15849	(91)	68M10	91A80	1	1	1	=	-	-		

Outline

1 Introduction

2 Example

3 Methods

4 Results

5 Discussion

6 Conclusions & Future

Matching Classifications

- 34 coincidences in 56 cases:
LLM primary classifications among arXiv classifications
- 17 cases:
one or more of arXiv primaries missing from LLM primary or secondary
- 1 apparent typo ignored

Matching Classifications — Example

- *The Kepler Cone, Maclaurin Duality and Jacobi-Maupertuis metrics* by Richard Montgomery.
- Primary 53: arXiv 2312.08173 [Math.DS] (70F; 53) 70G65 83C10
 - 70F Dynamics of a system of particles, including celestial mechanics
 - 70G65 Symmetries, Lie group and Lie algebra methods for problems in mechanics
 - 83C10 Equations of motion in general relativity and gravitational theory

MSC search key	arXiv ID	arXiv MSC	LLM primary	LLM secondary	# arXiv MSC	# LLM primary MSC	# primary missed	# secondary extra
53	2312.08173	(70F; 53)	70G65	83C10	2	1	1	1

Matching Classifications - Example

The Kepler Cone, Maclaurin Duality and Jacobi-Maupertuis metrics

Abstract. The Kepler problem is the special case $\alpha = 1$ of the power law problem: to solve Newton's equations for a central force whose potential is of the form μ/r^α where μ is a coupling constant. Associated to such a problem is a two-dimensional cone with cone angle $2\pi c$ with $c = 1/\alpha$. We construct a transformation taking the geodesics of this cone to the zero energy solutions of the α -power law problem. The 'Kepler Cone' is the cone associated to the Kepler problem. This zero-energy cone transformation is a special case of a transformation discovered by Maclaurin in the 1740s transforming the α -power law problem for any energies to a 'Maclaurin dual' γ -power law problem where $\gamma = \frac{2\alpha}{2\alpha - 1}$ and which, in the process, mixes up the energy of one problem with the coupling constant of the other. We derive Maclaurin duality using the Jacobi-Maupertuis metric reformulation of mechanics.

Matching Classifications — Example

- Submitted as DS (Dynamical Systems) and it's about the geometrical aspects of Newtonian mechanics and related matters.
- Intended to be fairly expository.
- arXiv gives 70F, which is fair enough, but arguably 70G35 is better, so we'll log agreement.
- In fact the relativity mention seems spurious, though Rutherford scattering is also discussed.
- There is a significant historical aside here too.

Differing Classifications

- 22 non-coincidences in 56 cases:
LLM primary classifications not among arXiv classifications
- 18 cases:
one or more of arXiv primaries missing from LLM primary or secondary
- 13 cases:
1 LLM additional secondary not among arXiv classes

Differing Classifications – Example

- *Symmetries and invariant solutions of the wave equation for shear disturbances in soft solids*
- arXiv:2303.01437 under 22, 76; LLM: 35Q72 74B20
 - 22 Topological groups, Lie groups For transformation groups, see 54H15, 57Sxx, 58-XX; for abstract harmonic analysis, see 43-XX
 - 76 Fluid mechanics For general continuum mechanics, see 74Axx, or other parts of 74-XX
 - 35Q Partial differential equations of mathematical physics and other areas of application [See also 35J05, 35J10, 35K05, 35L05
 - 35Q72 does not exist
 - 35Q70 PDEs in connection with mechanics of particles and systems of particles
 - 35Q74 PDEs in connection with mechanics of deformable solids
 - 74B20 Nonlinear elasticity

Differing Classifications – Example

- Here we have a paper that's given the pure mathematical classification of Lie theory and the applied classification of fluid mechanics.
- It is about PDE solutions though a Lie-theoretic approach is undertaken.

MSC search key	arXiv ID	arXiv MSC	LLM Primary	LLM Secondary	# arXiv MSC	# LLM primary MSC	# LLM primary "wrong"	LLM quality	# primary missed	# secondary extra
22	2303.01437	(22; 76)	35Q72	74B20	2	1	1	+2	2	1

- The application is in solids rather than fluids. So the LLM chatbot is better. Look at the abstract:

Differing Classifications – Example

Symmetries and invariant solutions of the wave equation for shear disturbances in soft solids

Abstract. The Lie-group approach was applied to determine symmetries of the third-order non-linear equation formulated for description of shear elastic disturbances in soft solids. Invariant solutions to this equation are derived and it turned out that they could represent outgoing or incoming exponentially decaying or unbounded disturbances.

Differing Classifications – Example

- MSC 35Q is correct and better
- 76 for fluids does not seem better than 74B for elastic materials (in this case soft solids).
- Some sort of 22 for the Lie-group approach seems good as a secondary.
- But a caveat in all this is that **35Q72 is a chatbot hallucination**: there is no such code in MSC2020!

Differing Classifications – How Bad?

- For each of the differing cases, we compared the LLM classification to the arXiv classification.
- Scored according to the following rubric:
 - +2 LLM better than arXiv class
 - +1 LLM slightly better than arXiv class
 - = arguable either way
 - 1 LLM slightly off
 - 2 LLM way off
- The proceedings paper gives a detailed analysis of each differing case.

Differing Classifications – How Bad?

- **Bottom Line:** In all but one of the differing cases the LLM classification was **as good or better** than the arXiv classification.
- Most often it was distinctly better.
- In the one bad case, it was way off.
- Two clear hallucinations – show stoppers for math.

Caveat Classifier

- Our experiment so far has a small sample size.
- Even then, we have a couple of cases of confabulation or hallucination.
- Trained neural nets are just prone to breakdown on outliers *cf.* Gary Marcus and Teslas hitting semi-trucks
- MSC classifying can be idiosyncratic when done by humans

Outline

- 1 Introduction
- 2 Example
- 3 Methods
- 4 Results
- 5 Discussion
- 6 Conclusions & Future**

Conclusions

- This is a very first step using an “off the shelf” model and a very small test sample.
- Even so, the results are very promising.
- The results are almost always as good or better than arXiv classification.

Next Steps

- Automate this interaction and examine a serious corpus.
 - Study quality of second-level classification.
 - Also do for ACM Classification.
-
- There are many alternative LLMs: Gemini, LLama, Mistral, ...
 - Train our own and see what we can do.
-
- Can we eliminate hallucinations?
 - Can we have fewer missed classifications?
 - What accuracy can we get with the whole paper as input?

Conclusions II

- The results are promising and merit further study.
- What does this say about the panorama and inter-relatedness of mathematics as a whole?
- Do the essential nuances of mathematical disciplines boil down to some modest size set of parameters?
- Can we extract a new and objectively better mathematical classification from a trained model?
- What can we learn about how mathematics evolves over time?