

CodeML Open RAIL-M v0.1 License

dated 22 December 2022

Section I: PREAMBLE

This OpenRAIL-M License was created under BigCode, an open and collaborative research project aimed at the responsible development and use of Large Language Models (“LLMs”) for code generation.

This license is generally applicable to any machine-learning Model used in the context of coding-related tasks. As with any other model, models applied to the context of coding tasks, such as code-generating models, have the potential to generate a broad positive impact in society and, at the same time, negative effects if misused.

This License strives for both the open and responsible use of the accompanying Model. When it comes to the open character, the License is inspired by open source permissive licenses for the grant of IP rights. Referring to responsible use, the addition of use restrictions not permitting the use of the Model in specific scenarios enables the Licensor to enforce the License in case potential misuse of the Model may occur. Even though derivative versions of the Model could be released under different licensing terms, the License specifies the obligation to include - at minimum - the same use restrictions as the ones in the original License (**this license**).

This License governs the use of the Model and its derivatives and is informed by the model card associated with the Model.

NOW THEREFORE, You and Licensor agree as follows:

1. Definitions

- (a) **"License"** means the terms and conditions for use, reproduction, and Distribution as defined in this document.
- (b) **"Data"** means a collection of information and/or content extracted from the dataset used with the Model, including to train, pretrain, or otherwise evaluate the Model. The Data is not licensed under this License.
- (c) **"Output"** means the results of operating a Model as embodied in informational content resulting therefrom.
- (d) **"Model"** means any accompanying machine-learning based assemblies (including checkpoints), consisting of learnt weights, parameters (including optimizer states), corresponding to the model architecture as embodied in the Complementary Material, that have been trained or tuned, in whole or in part on the Data, using the Complementary Material.

- (e) **“Derivatives of the Model”** means all modifications to the Model, works based on the Model, or any other model which is created or initialized by transfer of patterns of the weights, parameters, activations or output of the Model, to the other model, in order to cause the other model to perform similarly to the Model, including - but not limited to - distillation methods entailing the use of intermediate data representations or methods based on the generation of synthetic data by the Model for training the other model.
- (f) **“Complementary Material”** means the accompanying source code and scripts used to define, run, load, benchmark or evaluate the Model, and used to prepare data for training or evaluation, if any. This includes any accompanying documentation, tutorials, examples, etc, if any. Complementary Material is not licensed under this License.
- (g) **“Distribution”** means any transmission, reproduction, publication or other sharing of the Model or Derivatives of the Model to a third party, including providing the Model as a hosted service made available by electronic or other remote means - e.g. API-based or web-access.
- (h) **“Licensor”** means the rights owner or entity authorized by the rights owner that is granting the License, including the persons or entities that may have rights in the Model and/or distributing the Model.
- (i) **"You"** (or **"Your"**) means an individual or Legal Entity exercising permissions granted by this License and/or making use of the Model for whichever purpose and in any field of use, including usage of the Model in an end-use application - e.g. chatbot, translator, image generator.
- (j) **“Third Parties”** means individuals or legal entities that are not under common control with Licensor or You.
- (k) **"Contribution"** means any work, including the original version of the Model and any modifications or additions to that Model or Derivatives of the Model thereof, that is intentionally submitted to Licensor for inclusion in the Model by the rights owner or by an individual or Legal Entity authorized to submit on behalf of the rights owner. For the purposes of this definition, **“submitted”** means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Model, but excluding communication that is conspicuously marked or otherwise designated in writing by the rights owner as **"Not a Contribution."**
- (l) **"Contributor"** means Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Model.

Section II: INTELLECTUAL PROPERTY RIGHTS

Both copyright and patent grants apply to the Model and Derivatives of the Model. The Model and Derivatives of the Model are subject to additional terms as described in Section III, which shall govern the use of the Model and Derivatives of the Model even in the event Section II is held unenforceable.

2. Grant of Copyright License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare, publicly display, publicly perform, sublicense, and distribute the Model and Derivatives of the Model.

3. Grant of Patent License. Subject to the terms and conditions of this License and where and as applicable, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this paragraph) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Model and/or Derivatives of the Model where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Model or Derivatives of the Model to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Model or Derivative of the Model and/or a Contribution incorporated within the Model or Derivative of the Model constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for the Model and/or Derivative of the Model shall terminate as of the date such litigation is asserted or filed.

Section III: CONDITIONS OF USAGE, DISTRIBUTION AND REDISTRIBUTION

4. Distribution and Redistribution. You may host for Third Party remote access purposes (e.g. software-as-a-service), reproduce and distribute copies of the Model or Derivatives of the Model thereof in any medium, with or without modifications, provided that You meet the following conditions:

- a. Use-based restrictions as referenced in paragraph 5 **MUST** be included as an enforceable provision by You in any type of legal agreement (e.g. a license) governing the use and/or distribution of the Model or Derivatives of the Model, and You shall give notice to subsequent users You Distribute to, that the Model or Derivatives of the Model are subject to paragraph 5.
- b. You must give any Third Party recipients of the Model or Derivatives of the Model a copy of this License;
- c. You must cause any modified files to carry prominent notices stating that You changed the files;
- d. You must retain all copyright, patent, trademark, and attribution notices excluding those notices that do not pertain to any part of the Model, Derivatives of the Model.

You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions - **respecting** paragraph **4.a.** - for use, reproduction, or Distribution of Your modifications, or for any such Derivatives of the Model as a whole, provided Your use, reproduction, and Distribution of the Model otherwise complies with the conditions stated in this License.

5. Use-based restrictions. The restrictions set forth in Attachment A are considered Use-based restrictions. Therefore You cannot use the Model and the Derivatives of the Model for the specified restricted uses. You may use the Model subject to this License, including only for lawful purposes and in accordance with the License. Use may include creating any content with, finetuning, updating, running, training, evaluating and/or reparametrizing the Model. You shall require all of Your users who use the Model or a Derivative of the Model to comply with the terms of this paragraph (paragraph 5).

6. The Output You Generate. Except as set forth herein, Licensor claims no rights in the Output You generate using the Model. You are accountable for the Output you generate and its subsequent uses. No use of the output can contravene any provision as stated in the License.

Section IV: OTHER PROVISIONS

7. Updates and Runtime Restrictions. To the maximum extent permitted by law, Licensor reserves the right to restrict (remotely or otherwise) usage of the Model in violation of this License, update the Model through electronic means, or modify the Output of the Model based on updates.

8. Trademarks and related. Nothing in this License permits You to make use of Licensors' trademarks, trade names, logos or to otherwise suggest endorsement or misrepresent the relationship between the parties; and any rights not expressly granted herein are reserved by the Licensors.

9. Disclaimer of Warranty. Unless required by applicable law or agreed to in writing, Licensor provides the Model (and each Contributor provides its Contributions) on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied, including, without limitation, any warranties or conditions of TITLE, NON-INFRINGEMENT, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the appropriateness of using or redistributing the Model and Derivatives of the Model, and assume any risks associated with Your exercise of permissions under this License.

10. Limitation of Liability. In no event and under no legal theory, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (such as deliberate and grossly negligent acts) or agreed to in writing, shall any Contributor be liable to You for damages, including any direct, indirect, special, incidental, or consequential damages of any character arising as a result of this License or out of the use or inability to use the Model (including but not limited to damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses), even if such Contributor has been advised of the possibility of such damages.

11. Accepting Warranty or Additional Liability. While redistributing the Model, Derivatives of the Model thereof, You may choose to offer and charge a fee for, acceptance of support, warranty, indemnity, or other liability obligations and/or rights consistent with this License. However, in accepting such obligations, You may act only on Your own behalf and on Your sole responsibility, not on behalf of any other Contributor, and only if You agree to indemnify, defend, and hold each Contributor harmless for any liability incurred by, or claims asserted against, such Contributor by reason of your accepting any such warranty or additional liability.

12. If any provision of this License is held to be invalid, illegal or unenforceable, the remaining provisions shall be unaffected thereby and remain valid as if such provision had not been set forth herein.

END OF TERMS AND CONDITIONS

Attachment A

Use Restrictions

You agree not to use the Model or Derivatives of the Model:

- (a) In any way that violates any applicable national, federal, state, local or international law or regulation;
- (b) For the purpose of exploiting, harming or attempting to exploit or harm minors in any way;
- (c) To generate and/or disseminate malware (e.g. ransomware) or any other content to be used for the purpose of harming electronic systems;
- (d) To generate or disseminate verifiably false information and/or content with the purpose of harming others;
- (e) To generate or disseminate personal identifiable information that can be used to harm an individual;
- (f) To generate or disseminate information and/or content (e.g. images, code, posts, articles), and place the information and/or content in any public context (e.g. bot generating tweets) without expressly and intelligibly disclaiming that the information and/or content is machine generated;
- (g) To defame, disparage or otherwise harass others;
- (h) To impersonate or attempt to impersonate (e.g. deepfakes) others without their consent;
- (i) For fully automated decision making that adversely impacts an individual's legal rights or otherwise creates or modifies a binding, enforceable obligation;
- (j) For any use intended to or which has the effect of discriminating against or harming individuals or groups based on online or offline social behavior or known or predicted personal or personality characteristics;
- (k) To exploit any of the vulnerabilities of a specific group of persons based on their age, social, physical or mental characteristics, in order to materially distort the behavior of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm;
- (l) For any use intended to or which has the effect of discriminating against individuals or groups based on legally protected characteristics or categories;
- (m) To provide medical advice and medical results interpretation;
- (n) To generate or disseminate information for the purpose to be used for administration of justice, law enforcement, immigration or asylum processes, such as predicting an individual will commit fraud/crime commitment (e.g. by text profiling, drawing causal relationships between assertions made in documents, indiscriminate and arbitrarily-targeted use).