Intellectual Property Office

BL O/0657/23

04 July 2023

PATENTS ACT 1977

APPLICANT	Optum Inc.
ISSUE	Whether GB2114284.9 complies with Section 1(2) of the Patents Act 1977
HEARING OFFICER	Stephen Brown

DECISION

Introduction

1 Patent Application GB2114284.9 was published GB2600251 on 27th April 2022. It claims a priority date of 16th October 2020 from the US application 63092563. Despite several rounds of correspondence, the applicant has been unable to convince the Examiner that the application is allowable under Section 1(2) of the Act. As a consequence, the applicant was offered a hearing before a senior officer of the IPO and has requested that a decision be made on the papers.

The Application

2 The application is a computer implemented method that uses a machine learning model to predict an inferred disease risk score for a patient. In operation, a prediction analysis system receives a request from a client computer. The request is processed using the machine learning model to infer a hybrid disease risk score which it then outputs as notification data. The notification data comprises a patient name, a disease, the inferred or predicted risk score and some explanatory notes.

3 The generation of the machine learning model is described in further detail. The model itself is a graph based model and is generated using genetic programming operations which are in turn based on prior patient data objects. These genetic programming operations are used to generate a risk model based on further analysis that may include clinical prediction models and exiting outline clinical modules. In summary, the model is effectively trained using patient data, clinical data and existing models which can then be used to predict a disease risk for a patient.

Basis of this Decision

The Claims

4 To avoid any doubt this decision is based on the claims as filed on 16th January 2023. The independent claims of this set are attached as Annex A to this decision.

5 I have also attached a copy of the independent claims of an auxiliary set of claims also filed on 16th January at Annex B

The Applicant's Response

6 The last response from the applicant was received on 16th January 2023 in response to an earlier examination report issued by the Examiner on 16th Nov 2022. In that report the Examiner offered the applicant a hearing. In response, the applicant took up that offer and requested a hearing if the Examiner intended to refuse the application. Subsequent correspondence has led to the situation where the decision is required on the papers.

7 The applicant has filed no further arguments since their initial response when they filed amended and auxiliary claim sets. I will take this response as setting out the applicant's position and arguments.

The Law

8 The section of the Act concerning inventions excluded from patentability is Section 1(2). This reads:

> "It is hereby declared that the following (among other things) are not inventions for the purposes of this Act, that is to say, anything which consists of –
>
> ...
>
> (c) a scheme, rule or method for performing a mental act, playing a game or doing business or **a program for a computer**;
>
> …
>
> but the foregoing provision shall prevent anything from being treated as an invention for the purposes of this Act only to the extent that a patent or application for a patent relates to that thing as such."

9 In order to decide whether an invention relates to subject matter excluded by Section 1(2), the Court of Appeal has said that the issue must be decided by answering the question of whether the invention reveals a technical contribution to the state of the

art. The Court of Appeal in *Aerotel/Macrossan*[1] set out the following four-step approach to help decide the issue:

1) Properly construe the claim;

2) Identify the actual (or alleged) contribution;

3) Ask whether it falls solely within the excluded subject matter;

4) Check whether the actual or alleged contribution is actually technical in nature.

10 The operation of the approach is explained at paragraphs 40-48 of the judgment. Paragraph 43 confirms that identification of the contribution is essentially a matter of determining what it is the inventor has really added to human knowledge, and involves looking at substance, not form. Paragraph 47 adds that a contribution which consists solely of excluded matter will not count as a technical contribution.

11 The case law on computer implemented inventions has been further elaborated in *AT&T/CVON*[2] which provided five helpful signposts to apply when considering whether a computer program makes a relevant technical contribution. In *HTC v Apple*[3], Lewison LJ reconsidered the fourth of these signposts and felt that it had been expressed too restrictively. The revised signposts are:

i) whether the claimed technical effect has a technical effect on a process which is carried on outside the computer;

ii) whether the claimed technical effect operates at the level of the architecture of the computer; that is to say whether the effect is produced irrespective of the data being processed or the applications being run;

iii) whether the claimed technical effect results in the computer being made to operate in a new way;

iv) whether the program make the computer a better computer in the sense of running more efficiently and effectively as a computer; and

v) whether the perceived problem is overcome by the claimed invention as opposed to merely being circumvented.

12 As this application concerns the use of an AI in the form of machine learning, I have also referred to the guidelines issued by the IPO for "Examining patent applications relating to artificial intelligence (AI) inventions"[4]. Whilst I am aware that this does not

---

[1] Aerotel Ltd v Telco Holdings Ltd (and others) and Macrossan's Application [2006] EWCA Civ 1371

[2] AT&T Knowledge Ventures LP and CVON Innovations Limited v Comptroller General of Patents [2009] EWHC 343

[3] HTC v Apple [2013] EWCA Civ 451

[4] Examining patent applications relating to artificial intelligence (AI) inventions - GOV.UK (www.gov.uk)

supersede the law as it stands, it does provide a useful summary of the law as it currently is in relation to AI related inventions.

Step 1 – Construing the claims

13 The first step of the Aerotel test is to construe the claim. For brevity, my analysis will focus on claim 1 of the main set. However, as the other independent claims of this set have substantially the same scope, any conclusions will apply equally to them.

14 My understanding of claim 1 is that it relates to two aspects of the claimed invention. Specifically, there is an overarching method of predicting a disease risk score using a machine learning model. Within that umbrella, there is a set of steps that describe how the model is created so that it can form part of a prediction system to predict a disease risk. In essence, the claim provides for a prediction using a specially constructed machine learning model.

15 In understanding the claim, it is helpful to consider the meaning of the term “hybrid disease risk score” which is intrinsic to the purpose and scope of the claimed invention. Helpfully, paragraph 0023 provides a helpful view of the meaning of the term. That describes the term as a “risk score generated by a trained hybrid graph-based machine learning model by processing a set of graph-based feature embeddings for a corresponding patient data object”. My understanding of this is that it is effectively the prediction of the system. I feel supported in this by the discussion of Fig 8 at paragraph 0082 where the “inferred risk score* (801) for a particular disease is output to the screen. I note that fig 8 also helpfully explains the notification data at the end of the claim.

16 The remaining features of the claim listed as i – iv are the steps in the process of establishing the machine learning model that process the patient data in order to provide the risk score that is notified to the client or user machine as in Fig 8. The definitions of these features are helpfully listed in the specification between paragraph 0019 and paragraph 0024. In view of this, I see no need to explore these further.

17 There is one important point to make here. The steps in creating the machine learning model although framed in terms of graph based features relies entirely on patient or prior patient data objects, clinical predictions and a clinical risk model.

Product by Process

18 At this stage, and before I go onto to discuss the contribution made by the claim, I do need to consider the claim as a whole and whether it is a “product by process claim”. This is an objection raised by the examiner in their final report though, unfortunately, I do not have the benefit of the applicants’ views.

19 I have already stated that I see the claimed invention as a prediction made by a specially constructed machine learning model. It is without doubt a product (the prediction) that is arrived at by a process. Faced with the screen in Fig 8, I am none

the wiser as to how that has been arrived at as the end user. I know I have submitted a query to the system, but I have no view on how the results have been arrived at. In reality, I am likely unaware that it has been arrived at by machine learning.

20 There is however, one slight difference from a standard, if there is such a thing, product by process claim. Specifically, the construction of the machine learning model is not strictly a method of manufacture but more a matter of a way of configuring the model to achieve a specific result. The question, I therefore need to address is whether, or not, the steps taken in construction of the machine learning model are the equivalent of a method of manufacture.

21 In the circumstances of this application, as it is currently claimed, I believe they are one in the same. The claim tells me, how the machine learning model is trained using a series of genetic programming operations within a graph based model. It is to all intents and purposes analogous to the situation in paragraph 14.120.1 of the Manual of Patent Practice when discussing *Kirin Amgen Inc v Hoechst Marion Roussel*[5]. That paragraph relates an example of an apparatus comprising the features A and B where feature A is treated by an oven B. It is not dissimilar, to the situation of the claim where a data query (A) is treated by the machine learning model (the oven "B"). In view of the similarity and that I have no wish to divert form the precedent set by the House of Lords in *Kirin Amgen*[5], there is only one conclusion I can come to, and that is that claims 1, 10 and 18 as they stand are "product by process" claims. As such, I must treat the resultant, the prediction, as the product.

Step 2- Identifying the Contribution

22 The next step is to identify the contribution made by the claimed invention. It is perhaps helpful at this point to point out that the examiner and attorney do not entirely agree on this.

23 The attorney has identified as the contribution as:

> *"The provision of a hybrid disease risk score generated using a set of genetic programming techniques to an end-user in response to that user sending a query to the predictive data analysis system for a patient data object, where the provision includes generating notification data as part of a set of prediction-based actions which are performed based on a hybrid disease risk score.*
>
> *The hybrid disease risk score is thus a risk score which is in effect dynamically calculated responsive to the received query for that patient data object using an inferred HDRSGML model having the features recited in claim 1. The risk score provides better accuracy and reliability as it is generated using a set of genetic programming operations, for example, a set of symbolic regression operations (see claim 8), which generate a risk model for a graph-based feature embedding data object based at least in part on (a) symbolic*

[5] *Kirin Amgen Inc v Hoechst Marion Roussel Ltd [2005] RPC 9*.

regression, (b) initial clinical prediction models, and (c) an existing outline clinical risk model for a disease associated with the hybrid disease risk score."

24 The examiner on the other hand sees the contribution as:

> *"a computer program to, responsive to receiving a query, process graph feature embedding data objects for a patient data object using an inferred hybrid risk score generation machine learning model to generate a hybrid risk score, and perform prediction-based actions based on the hybrid risk score, including generating and providing a notification.*
>
> *The inferred hybrid risk score generation machine learning model is generated using genetic programming operations based on inferred hybrid risk scores for prior patient data objects generated by a hybrid graph-based machine learning model"*

25 To some extent identifying the contribution relies on my view of the claim as a "product by process" . Specifically, does the contribution include the process (steps i-iv)) by which the prediction is arrived at or not? In order to make progress from here, I am going to set out two views of the contribution. The first is to treat the contribution as a product and the second to consider the underlying machine learning model in case I am wrong about my conclusion on product by process.

26 Taking the first of these, I note that both the examiner and the applicant are close to agreement as to notifying a prediction to a user. The applicant differs slightly in that they consider the use of genetic programming techniques as part of this contribution.

27 For my part, there is no suggestion of any new computing equipment. I also take the view that the idea of imputing information at a client machine, processing it in a system and providing an output to a user interface on the client machine is so well known that it cannot form part of the contribution.

28 Turning to the applicant's inclusion of the "genetic programming techniques". I believe this is covered by the views I have expressed on "product by process" and as such it is not relevant to the first contribution.

29 It follows that the first contribution I have identified is:

> *"a computer implemented method of predicting a disease risk for a patent by analysing the patient data (using a machine learning model) to provide a notification of the risk score."*

30 Turning now to the second contribution, I need to consider the underlying machine learning model as if the claim was not product by process. That is, I need to consider the underlying machine learning model as part of the contribution.

31 The applicant argues that the contribution made by the graph based machine learning model is more accurate and reliable because of the use of genetic programming operations. The examiner, on the other hand, excludes the accuracy and reliability as part of their contribution.

32 I do not see how greater accuracy or reliability can be specifically attributed to using a graph-based system. I understand that using graph based system in a machine learning environment can form the basis of a supervised learning system. However, that is more to do with the mechanics of machine learning and allows it to use unlabelled data as opposed to a more traditional unsupervised system that relies on labelled data.

33 I offer one further observation here. In relying on the "graph based nature" of the machine learning model, I am being invited to take the view that this is a determinative factor in whether the application may be allowable. In effect, I am being asked to indulge in form over substance and it is not an avenue I wish to go down. I am comfortable with the view that, in substance, this application is the use of machine learning to identify a disease risk score. The type of machine learning used is a matter of form in this case. It may well be a novel use but that is not the issue here.

34 Consequently, I take the view that the Examiners version of the contribution is the more accurate. I thus view the second, or alternative, contribution as:

> *"a computer implemented method of predicting a disease risk for a patent by analysing the patient data using a hybrid risk score generation machine learning model to provide a notification of the risk score wherein the inferred risk is generated using genetic programming operations based on inferred hybrid risk scores for prior patient data objects generated by a machine learning model"*

Step 3 – Does it falls solely within excluded subject matter?

35 The third step in the Aerotel test is to consider if the contribution lies solely in an excluded area. Both contributions I have identified are computer programs. In view of that I need to decide whether these contributions comprise relevant technical contributions.

The First Contribution

36 The first contribution is a method of predicting a disease risk using a machine learning model. In terms of the recent guidance issued by the IPO on AI[4] this is clearly a case of an "applied AI". This leads me to the questions set out in the HTC[3] signposts (i), (ii) and (v).

37 Taking (v) first, what is the problem being solved? I have no doubt that there is clearly a problem in identifying a risk of disease, but I question if it is technical beyond that it is a prediction. It does not solve a problem although it may inform the user. As such, I do not consider that the contribution meets signpost (v).

38 Turning to signpost (i), the applicant has argued that notifying the result to a screen counts as part of the contribution and hence signpost 1 is met. However, this is nothing more than one would expect of any computer system. For that reason, I do not consider it part of the contribution. Consequently, I can see no control of anything beyond the boundaries of the computer system and signpost (i) is not met.

39 Turning to signpost (ii), it is clear that the data being input is patient data, this is processed against a model containing collated patient data and clinical data relating to the disease to output a prediction based on a comparison of the two. It further follows that the method is entirely reliant of the data being processed. Thus the contribution does not meet signpost (ii).

40 For completeness, I will comment that since the computer, including the machine learning module, appears to operate in a completely standard way, the contribution cannot result in the computer itself operating in a new way, or running more efficiently or effectively. Thus, this contribution does not meet signposts (iii) or (iv) either.

41 I therefore conclude that the first contribution is excluded as a computer program, as such.

The Second Contribution

42 The second contribution raises the question of whether the use of "genetic programming techniques" are enough to provide a non-excluded technical contribution. I have already discussed signposts (i), (ii) and (v) above and my view on these is not changed by the additional subject matter in this contribution. Neither do I consider signpost (iii) as offering any help – it is still not a new computer, nor does it operate in a different way.

43 I do, however, have to give some consideration to signpost (iv) and consider whether the computer operates more efficiently or effectively as a computer. The claim clearly uses a set of programming techniques and I need to ask whether this makes the computer more effective or efficient. I need only to look at the case of *Autonomy Corp*[6] for guidance. I note that the Examiner has already raised this in their latest report quoting paragraph 29, part (viii), namely:

> *"The mere fact that a computer program reduces the load on the processor or makes economical use of the computer's memory or makes more efficient use of the computer's resources does not amount to making a better computer, and thus does not take it outside the category of computer program as such."*

44 This makes it clear to me that a mere programming technique is not enough. As a result, I believe the second contribution is also excluded as a program for a computer as such.

Step 4 - Is the contribution technical in nature?

45 Since I have decided that both contributions do not have a technical effect beyond that of a program running on a computer, they also fail this step of the test. I thus decide that the invention, as defined in the main claims, is excluded under section 1(2).

---

[6] *Autonomy Corp Ltd v Comptroller General of Patents, Trade Marks and Designs [2008] EWHC 146 (Pat)*

The Auxiliary Request

46 The differences in the auxiliary request made the attorney can be summarised as:

a. Incorporating a specific reference to notifying a risk score to an end user via a user communications network: and

b. Generating user interface data to be displayed on the end user computer.

47 Referring to the additional reference to a communications network. There is no doubt in my mind that such networks are now so commonplace that the use of one would not form part of the contribution. If anything, it is a statement of the obvious and consequently does not help the applicant.

48 Similarly, sending data to a user machine causing it to display the data does not add anything to the contribution. It is what computers do and there is no suggestion that this adds anything technical.

49 On the basis that neither of the changes alter either of the contributions I have identified earlier in this decision, I conclude that the auxiliary claims are also excluded under section 1(2).

Conclusion

50 For the reasons given above, I decide that the invention as set out in the claims is excluded under Section 1(2)(c) as a program for a computer, as such. Having reviewed the application, I do not consider that any saving amendments are possible. I therefore refuse the application under Section 18(3).

Appeal

51 Any appeal must be lodged within 28 days after the date of this decision.

Dr Stephen Brown

Deputy Director, acting for the Comptroller

Annex A – Main claims

Claim 1

A Computer-implemented method of providing a hybrid disease risk score for a patient data object to an end user, wherein the hybrid disease risk score is based at least in part on at least one of patient genomic data, patient behavioural data or patient clinical data, the computer implemented method comprising:

At a predictive data analysis system configured to receive predictive data analysis requests from a plurality of client computing entities, responsive to receiving a query for a patient data object via a user interface or a client computer entity of an end user:

processing by one or more processors, a plurality of graph based feature embedding data objects for the patient data object using an inferred hybrid disease risk score generation machine learning model to generate the hybrid disease risk score, wherein a graph based feature embedding data object comprises a vector of one or more values that are determined based at least in part on a risk tensor of one or more tensors for the patient data object; wherein:

(i) the inferred hybrid disease risk score generation machine learning model comprises an inferred hybrid disease risk score generation machine learning model generated using a set of genetic programming operations:

(ii) the set of genetic programming operations comprises a set of genetic programming operations performed based at least in part on a set of inferred hybrid disease risk scores for a set of prior patient data objects'

(iii) the set of genetic programming operations comprises a set of genetic programming operations which generate a risk model for a graph based feature embedding data object at least in part on (a) symbolic regression, (b) initial clinical prediction models, and (c) an existing outline clinical risk model for a disease associated with the hybrid risk score, and

(iv) each inferred hybrid disease risk score of the set of inferred hybrid disease risk scores is an inferred hybrid disease risk score generated by processing a plurality of prior graph based feature embedding data objects for a corresponding prior patient data object of the set of prior patient data objects using a trained hybrid graph based machine learning model, and

Performing using one of more processors, one or more prediction based actions based at least in part on the hybrid disease risk score, wherein the prediction based actions comprise:

Generating notification data for one or more notification user interfaces; and providing the notification data to the client computing entity of the end user;

Wherein the notification data comprises an indication of the patient, a disease, the inferred hybrid risk disease risk score, and explanatory metadata for the inferred hybrid disease risk score.

Claim 10

An apparatus for providing a hybrid disease risk score for a patient data object to an end-user, wherein the hybrid disease risk score is based at least in part on at least one of patient genomic data, patient behavioural data, or patient clinical data, the apparatus comprising at least on processor and at least one memory including program code, the at least one memory and the program code configured to, with the processor cause the apparatus to at least:

responsive to receiving a query for a patient data object via a user interface of a client computer entity of an end user:

process, by one or more processors, a plurality of graph-based feature embedding data objects for the patient data object using an inferred hybrid disease risk score generation machine learning model to generate the hybrid disease risk score, wherein a graph-based feature embedding data object comprises a vector of one or more values that are determined based at least in part on a risk tensor of one or more risk tensors for the patient data object, wherein:

(i) the inferred hybrid disease risk score generation machine learning model comprises an inferred hybrid disease risk score generation machine learning model generated using a set of genetic programming operation,

(ii) the set of genetic programming operations comprises a set of genetic programming operations performed based at least in part on a set of inferred hybrid disease risk scores for a set of prior patient data objects,

(iii) the set of genetic programming operations comprises a set of genetic programming operations which generate a risk model for a graph-based feature embedding data object based at leas in part on (a) symbolic regression, (b) initial clinical prediction methods, and (c) an existing outline clinical risk model for a disease associated with the hybrid disease risk score, and

(iv) each inferred hybrid disease risk score of the set of inferred hybrid disease risk scores is an inferred hybrid disease risk score generated by processing a plurality of prior graph-based feature embedding data objects for a corresponding prior patient data object of the set of prior patient data objects using a trained hybrid graph-based machine learning model: and

perform using the one or more processors one or more prediction-based actions based at least in part on the hybrid disease risk score wherein the prediction based actions comprise:

generating notification data for one or more notification user interfaces:

and

providing the notification data to the client computing entity of the end user, wherein the notification data comprises an indication of the patient, a disease, the inferred hybrid disease risk score, and explanatory metadata for the inferred hybrid disease risk score, where in the apparatus comprises a predictive analysis system configured to receive predictive data analysis requests from a plurality of client computing entities.

Claim 18

A computer program product for providing a hybrid disease risk score for a patient data object to an end-user, wherein the hybrid disease risk score is based at least in part on at least one of patient genomic data, patient behavioural data, or patient clinical data, the computer program product comprising at least one non-transitory computer readable storage medium having computer readable program code portions stored thereon, the computer-readable program code portions configured to:

Cause at a predictive data analysis system configured to receive predictive data analysis requests from a plurality of client computing entities, responsive to receiving a query for a patient data object via a user interface of a client computer entity of an end user, the predictive analysis system to:

process, by one or more processors, a plurality of graph-based feature embedding data objects for the patient data object using an inferred hybrid disease risk score generation machine learning model to generate the hybrid disease risk score, wherein a graph-based feature embedding data object comprises a vector of one or more values that are determined based at least in part on a risk tensor of one or more risk tensors for the patient data object, wherein:

(v) the inferred hybrid disease risk score generation machine learning model comprises an inferred hybrid disease risk score generation machine learning model generated using a set of genetic programming operation,

(vi) the set of genetic programming operations comprises a set of genetic programming operations performed based at least in part on a set of inferred hybrid disease risk scores for a set of prior patient data objects,

(vii) the set of genetic programming operations comprises a set of genetic programming operations which generate a risk model for a graph-based feature embedding data object based at least in part on (a) symbolic regression, (b) initial clinical prediction methods, and (c) an existing outline clinical risk model for a disease associated with the hybrid disease risk score, and

(viii) each inferred hybrid disease risk score of the set of inferred hybrid disease risk scores is an inferred hybrid disease risk score generated by processing a plurality of prior graph-based feature embedding data objects for a corresponding prior patient data object of the set of prior patient data objects using a trained hybrid graph-based machine learning model: and

perform using the one or more processors one or more prediction-based actions based at least in part on the hybrid disease risk score wherein the prediction based actions comprise:

generating notification data for one or more notification user interfaces:

and

providing the notification data to the client computing entity of the end user,

wherein the notification data comprises an indication of the patient, a disease, the inferred hybrid disease risk score, and explanatory metadata for the inferred hybrid disease risk score.

Annex B – Auxiliary Claims

Claim 1

A Computer-implemented method of *electronically notifying* ~~providing~~ a hybrid disease risk score for a patient data object to an end user *via at least one communications network*, wherein the hybrid disease risk score is based at least in part on at least one of patient genomic data, patient behavioural data or patient clinical data, the computer implemented method comprising:

at a predictive data analysis system configured to receive predictive data analysis requests from a plurality of client computing entities, responsive to receiving a query *via at least one communications network*, for a patient data object via a user interface or a client computer entity of an end user:

processing by one or more processors, a plurality of graph based feature embedding data objects for the patient data object using an inferred hybrid disease risk score generation machine learning model to generate the hybrid disease risk score, wherein a graph based feature embedding data object comprises a vector of one or more values that are determined based at least in part on a risk tensor of one or more tensors for the patient data object; wherein:

(i) the inferred hybrid disease risk score generation machine learning model comprises an inferred hybrid disease risk score generation machine learning model generated using a set of genetic programming operations:

(ii) the set of genetic programming operations comprises a set of genetic programming operations performed based at least in part on a set of inferred hybrid disease risk scores for a set of prior patient data objects'

(iii) the set of genetic programming operations comprises a set of genetic programming operations which generate a risk model for a graph based feature embedding data object at least in part on (a) symbolic regression, (b) initial clinical prediction models, and (c) an existing outline clinical risk model for a disease associated with the hybrid risk score, and

(iv) each inferred hybrid disease risk score of the set of inferred hybrid disease risk scores is an inferred hybrid disease risk score generated by processing a plurality of prior graph based feature embedding data objects for a corresponding prior patient data object of the set of prior patient data objects using a trained hybrid graph based machine learning model, and

Performing using one of more processors, one or more prediction based actions based at least in part on the hybrid disease risk score, wherein the prediction based actions comprise:

Generating user interface data for one or more prediction output user interfaces and providing the user interface data to the client computing entity of the end user via the at least one communication network;

Causing the display of one or more prediction output interfaces to the client computing entity of the end user;

Generating notification data for one or more notification user interfaces; and providing the notification data to the client computing entity of the end user *via the at least one communications network; and*

Causing one or more electronically-generated notifications to be presented on the client computing entity of the end user, Wherein the notification data comprises an indication of the patient, a disease, the inferred hybrid risk disease risk score, and explanatory metadata for the inferred hybrid disease risk score.

Claim 10

An apparatus for *electronically notifying* ~~providing~~ a hybrid disease risk score for a patient data object to an end-user *via at least one communications network*, wherein the hybrid disease risk score is based at least in part on at least one of patient genomic data, patient behavioural data, or patient clinical data, the apparatus comprising at least on processor and at least one memory including program code, the at least one memory and the program code configured to, with the processor cause the apparatus to at least:

responsive to receiving a query *via at least one electronic communications network* for a patient data object via a user interface of a client computer entity of an end user:

process, by one or more processors, a plurality of graph-based feature embedding data objects for the patient data object using an inferred hybrid disease risk score generation machine learning model to generate the hybrid disease risk score, wherein a graph-based feature embedding data object comprises a vector of one or more values that are determined based at least in part on a risk tensor of one or more risk tensors for the patient data object, wherein:

(i) the inferred hybrid disease risk score generation machine learning model comprises an inferred hybrid disease risk score generation machine learning model generated using a set of genetic programming operation,

(ii) the set of genetic programming operations comprises a set of genetic programming operations performed based at least in part on a set of inferred hybrid disease risk scores for a set of prior patient data objects,

(iii) the set of genetic programming operations comprises a set of genetic programming operations which generate a risk model for a graph-based feature embedding data object based at least in part on (a) symbolic regression, (b) initial clinical prediction methods, and (c) an existing outline clinical risk model for a disease associated with the hybrid disease risk score, and

(iv) each inferred hybrid disease risk score of the set of inferred hybrid disease risk scores is an inferred hybrid disease risk score generated by processing a plurality of prior graph-based feature embedding data objects for a corresponding prior patient data object of the set of prior patient data objects using a trained hybrid graph-based machine learning model: and

perform using the one or more processors one or more prediction-based actions based at least in part on the hybrid disease risk score wherein the prediction based actions comprise:

generating user interface data for one or more prediction output user interfaces and providing the user interface data to the client computing entity of the end user via the at least one communications network;

causing display of the one or more prediction output user interfaces to the client computing entity of the end user;

generating notification data for one or more notification user interfaces:

and

providing the notification data to the client computing entity of the end user *via the at least one communication network and*

causing one or more electronically-generated notifications to be presented on the client computing entity of the end user, wherein the notification data comprises an indication of the patient, a disease, the inferred hybrid disease risk score, and explanatory metadata for the inferred hybrid disease risk score, where in the apparatus comprises a predictive analysis system configured to receive predictive data analysis requests from a plurality of client computing entities.

Claim 18

A computer program product for *electronically notifying* ~~providing~~ a hybrid disease risk score for a patient data object to an end-user *via at least one communications network*, wherein the hybrid disease risk score is based at least in part on at least one of patient genomic data, patient behavioural data, or patient clinical data, the computer program product comprising at least one non-transitory computer readable storage medium having computer readable program code portions stored thereon, the computer-readable program code portions configured to:

Cause at a predictive data analysis system configured to receive predictive data analysis requests from a plurality of client computing entities, responsive to receiving a query *via at least one communications network* for a patient data object via a user interface of a client computer entity of an end user, the predictive analysis system to:

process, by one or more processors, a plurality of graph-based feature embedding data objects for the patient data object using an inferred hybrid disease risk score generation machine learning model to generate the hybrid disease risk score, wherein a graph-based feature embedding data object comprises a vector of one or more values that are determined based at least in part on a risk tensor of one or more risk tensors for the patient data object, wherein:

(i) the inferred hybrid disease risk score generation machine learning model comprises an inferred hybrid disease risk score generation machine learning model generated using a set of genetic programming operation,

(ii) the set of genetic programming operations comprises a set of genetic programming operations performed based at least in part on a set of inferred hybrid disease risk scores for a set of prior patient data objects,

(iii) the set of genetic programming operations comprises a set of genetic programming operations which generate a risk model for a graph-based feature embedding data object based at least in part on (a) symbolic regression, (b) initial clinical prediction methods, and (c) an existing outline clinical risk model for a disease associated with the hybrid disease risk score, and

(iv) each inferred hybrid disease risk score of the set of inferred hybrid disease risk scores is an inferred hybrid disease risk score generated by processing a plurality of prior graph-based feature embedding data objects for a corresponding prior patient data object of the set of prior patient data objects using a trained hybrid graph-based machine learning model: and

perform using the one or more processors one or more prediction-based actions based at least in part on the hybrid disease risk score wherein the prediction based actions comprise:

generating user interface data for one or more prediction output user interfaces and providing the user interface data to the client computing entity of the end user via the at least one communications network;

causing display of the one or more prediction output user interfaces to the client computing entity of the end user;

generating notification data for one or more notification user interfaces:

and

providing the notification data to the client computing entity of the end user *via the at least one communications network; and*,

causing one or more electronically-generated notifications to be presented on the client computing entity of the end user, wherein the notification data comprises an indication of the patient, a disease, the inferred hybrid disease risk score, and explanatory metadata for the inferred hybrid disease risk score.