



## Smart Management Systems

Client: Departament d'Agricultura,  
Ramaderia, Pesca i Alimentació

**Activitat de demostració:**  
**Analítica avançada en granges de porcs**

**LL 1 - Document de bones pràctiques per a la monitorització de granges**



Generalitat de Catalunya  
**Departament d'Agricultura,  
Ramaderia, Pesca i Alimentació**



Fons Europeu Agrícola  
de Desenvolupament Rural:  
Europa inverteix en les zones rurals



**Informe redactat per Marc Ribalta, Lluís Echeverria i Josep Pijuan,**  
**Investigadors**  
**Validat per Xavier Domingo,**  
**Director de Unitat**

Títol:	Analítica avançada en granges de porcs: Bones pràctiques en la monitorització de granges
Oferta número:	56 30015 2018 2A
Responsable tècnic	Josep Pijuan Parra <a href="mailto:josep.pijuan@eurecat.org">josep.pijuan@eurecat.org</a> 973 193 660
Data:	30-09-2020
Client:	Departament d'Agricultura, Ramaderia, Pesca i Alimentació
Receptor:	<b>Servei d'Innovació Agroalimentària</b> Departament d'Agricultura, Ramaderia, Pesca i Alimentació Gran Via de les Corts Catalanes, 612-614   08007 Barcelona
Pàgines	10



## Índex

1.	INTRODUCCIÓ .....	4
2.	QUÈ CAL MONITORITZAR .....	5
3.	COM CAL FER-HO.....	7
4.	PERQUÈ CAL FER-HO.....	9

## 1. Introducció

En l'era de la ramaderia 4.0; la digitalització, la monitorització de processos i l'anàlisi aplicada a granges, són activitats fonamentals per garantir la millora continua dels processos productius, i el suport de les eines digitals s'està convertint en un requeriment crític per aconseguir-ho. En paral·lel, tecnologies com la Intel·ligència Artificial (IA) ens permeten utilitzar les dades generades per extreure'n coneixement i poder prendre així decisions intel·ligents, adequades a les circumstàncies reals de l'entorn i que assegurin el màxim rendiment (maximitzant el resultat mentre és minimitzen costos i altres aspectes claus com la contaminació o compliment de normatives) de la cadena de valor.

Les dades són l'element clau per al correcte desenvolupament de sistemes de suport a la decisió, basats en algorismes supervisats i no supervisats de l'àmbit de l'aprenentatge automàtic, compresos en el catàleg de tècniques de la IA. Aquests algorismes 'aprenen' de les dades històriques i apliquen aquest coneixement al present, essent capaços de predir comportaments futurs (manteniment predictiu, prediccions de qualitat) i recomanar configuracions de procés. Per tant, la correcta captura d'aquestes dades resulta crítica per a la generació de models intel·ligents efectius que permetin realitzar les accions més òptimes i adequades en cada cas. En termes generals, la quantitat i qualitat de les dades estan directament relacionades amb la qualitat final dels models intel·ligents.

La *monitorització*, pretén ser conscient de l'estat d'un sistema mitjançant l'observació dels canvis que s'hi poden produir en el temps. L'ús de sensors permet controlar aquests canvis mitjançant la mesura de magnituds físiques o químiques, que a la vegada poden ser transformades en dades susceptibles de ser emmagatzemades o utilitzades en processos de control. Totes les dades generades en un procés són la matèria prima per a la millora contínua d'aquest.

Aquest document presenta les experiències, lliçons apreses i conclusions extretes, per part de l'equip de la unitat de Smart Management Systems d'Eurecat com a experts en anàlisi de dades, en el projecte "Anàlisi avançada de dades en granges de porcs". L'esperit del projecte persegueix mostrar al sector ramader català els punts clau a seguir per tal de poder treure el màxim profit de la tecnologia actual en quan a monitorització i posterior anàlisi de dades. En el projecte s'ha treballat per millorar la productivitat de granges d'animals, específicament de truges reproductores. En aquest sentit, mitjançant tècniques de mineria de dades i Intel·ligència Artificial, i gràcies a la col·laboració amb l'empresa OPP Group, s'han analitzat i processat dades històriques en diversos cicles reproductius, recollides mitjançant el seu software de gestió, i s'han desenvolupat models intel·ligents amb l'objectiu de millorar el rendiment dels processos productius.

Partint de la base que la obtenció d'un conjunt de dades de qualitat és imprescindible per al treball amb tècniques d'aprenentatge automàtic, les següents seccions descriuen els punts més importants a tenir en compte per aconseguir-ho. Al llarg de la secció 2, es descriu quina informació s'hauria de monitoritzar i recollir per poder tenir una base de treball suficient que permeti dur a terme estudis d'aquesta tipologia i envergadura de forma satisfactòria. A continuació, la secció 3 explica com cal fer aquesta recollida d'informació, i quines especificacions s'han de seguir per aconseguir una estructura de dades fàcil de gestionar. Finalment, la secció 4 esmenta el raonament de les especificacions, perquè cal fer-ho i a quins objectius es pot arribar seguint una bona disciplina.

## 2. Què cal monitoritzar

En termes generals, i sempre tenint en compte l'objectiu final de l'estudi, la monitorització de granges es pot dividir en dos grans grups: captura de variables d'entorn i operacionals, i captura de informació de l'animal.

D'un mode similar a com les variables d'entorn d'un procés industrial caracteritzen i afecten a com aquest opera, la informació de l'entorn i la configuració de la granja afecta al seu dia a dia i caracteritza els seus processos fent que, per un mateix procés, granges o entorns diferents requereixin de configuracions diferents. Cada granja té, per exemple, unes condicions climàtiques específiques que al llarg del temps es converteixen en una història només vista per la "mateixa" granja. D'aquesta manera, si es desitja poder arribar a extrapolar i generalitzar coneixement d'una granja a una altra, per exemple, tota condició climàtica que pugui afectar la vida de l'animal, haurà de ser capturada. D'aquesta manera, és important monitoritzar elements com la temperatura ambient i la humitat dins i fora de la granja, així com altres condicions climàtiques exteriors, sempre lligant cada observació amb l'instant de captura de la informació.

A partir d'aquí, la captura d'informació a nivell d'animal constitueix la principal metodologia per adquirir les dades necessàries per a, posteriorment, poder estudiar el comportament d'aquests i les relacions causa-efecte entre els diferents esdeveniments al llarg de la seva vida, del cicle de treball i de la cadena de producció. Diferents estudis poden requerir de diferent tipus d'informació, diferents fonts, i diferents granularitats, de forma que resulta molt complicat identificar, a priori, quines variables seran les bones per cada cas d'estudi. Disposar del coneixement expert del sector sempre és un bon punt de partida, però en l'era de la Intel·ligència Artificial, els algoritmes intel·ligents poden identificar relacions crítiques en les dades que resulten clau per a la solució excel·lent de les diferents problemàtiques a resoldre.

A continuació és llista un conjunt de variables que s'ha observat, al llarg del projecte presentat i relacionat amb la gestió de truges reproductives, que aporten informació als models intel·ligents per a la solució de l'objectiu d'aquest:

- Origen: Origen de l'animal. Per conèixer el passat de l'animal i tenir-ne traçabilitat. Es pot extreure diferents conclusions i establir relacions a partir de la granja d'origen.
- Raça i altres variables identificadores del tipus: Permeten l'estudi i diferenciació d'animals.
- Pes: El pes de l'animal hauria de ser una mesura que es realitzi de forma diària per poder disposar de l'evolució al llarg de la seva vida. Aquesta variable és de vital importància, ja que identifica i representa gran quantitat d'anomalies i resulta clau per identificar possibles problemàtiques.
- Esdeveniments i dates: Diferents esdeveniments que han succeït als animals al llarg de la seva instància en la granja. És molt important identificar el moment exacte en que van succeir per creuar-ho amb la història productiva de l'animal:
  - o Data de naixement.
  - o Data d'arribada a la granja
  - o Data de baixa de la granja.

- Moviments d'animals: Moviments d'animals per les diferents zones de la granja. Això permetrà creuar informació espai-temporal (com la climàtica) d'aquestes zones amb els animals i poder identificar, per exemple, problemes en localitzacions específiques o grups d'animals.
  - Inseminacions: Data d'inseminacions, per embaràs, d'una truja. Permet estudiar la tendència de les inseminacions durant els diferents cicles de la truja i per avaluar el seu rendiment.
  - Parts: Data de part. Permet estudiar tendències i canvis de comportament i desgast.
  - Avortaments: Data d'avortament d'una truja. Si una truja avorta, pot canviar el patró de comportament d'aquesta en els futurs cicles.
  - Deslletaments
  - Patologies: Patologies i malalties. Les patologies indiquen un mal estat d'una truja i habitualment van lligades a anomalies en el procés. La data va lligada a una variable categòrica de tipus de malaltia.
  - Medicaments: Per cada patologia s'administra certs medicaments als animals. Guardar els medicaments administrats a analitzar la gravetat d'una patologia i la recuperació d'una truja i el seu posterior comportament. La data va lligada a una variable categòrica de tipus de medicament.
- Causa de la baixa: una truja pot causar baixa d'una granja per diversos motius om per exemple: baixa productivitat, baixa condició física, malaltia o mort.
  - Garrins nascuts: Nombre de garrins nascuts per part (lligat a un part). Gràcies a aquest nombre es pot saber l'evolució de la truja en tots els cicles.
  - Garrins vius: Nombre de garrins vius (lligat a un part). De les cries nascudes, és important saber el nombre de garrins vius per estudiar comportaments, evolucions i anomalies.
  - Garrins morts: Nombre de garrins morts (lligat a un part). Dels garrins nascuts morts, és important saber-ne el nombre per estudiar comportaments, evolucions i anomalies
  - Garrins deslletats: Nombre de garrins deslletats (per cicle). El nombre de garrins deslletats indica la capacitat lletera d'una truja. Conèixer l'evolució d'aquesta capacitat és important per l'estudi.
  - Alimentació: L'alimentació és habitualment un tipus de informació important a estudiar. És important, en la mesura del possible, recollir quin tipus de ingredients, i en quina quantitat, ha consumit l'animal, i en quina data, per lligar-ho amb la història. El mateix aplica a l'aigua.
  - Garrins adoptats: Nombre de garrins adoptats (per cicle), procedents d'una altra truja per a ser alimentats.
  - Garrins retirats: Nombre de garrins retirats (per cicle), cap a una altra truja per a ser alimentats.
  - Origen i destí de cries: Origen i destinació de les cries adoptades o retirades (per part). Per millorar l'anàlisi intern d'una granja, la indicació de l'origen i destí ens permet estudiar i utilitzar relacions entre diferents truges.
  - Condició física de truges i garrins: Aquesta informació és subjecte d'estudi en diversos projectes de recerca i si pot ser definida de forma objectiva pot permetre descriure millor el potencial productiu de les truges.

Del mateix mode que en les dades contextuals, aquesta informació sempre ha d'anar lligada de l'instant de mesura per poder, en temps d'anàlisi, crear i relacionar les dades.

Habitualment, la monitorització individualitzada d'algunes d'aquestes variables correspon a uns costos en equipament elevats. Variables com la quantitat d'aigua o l'aliment consumit per un animal requereixen d'una infraestructura avançada que habitualment es cara. Per evitar aquestes inversions és habitual capturar la informació de forma agregada, com per exemple litres d'aigua que s'han consumit al llarg del dia en una zona específica, i extrapolar els valors als animals que tenim identificats que és trobaven allí.

### 3. Com cal fer-ho

La recollida de la informació s'ha de fer seguint una sèrie de pautes i bones pràctiques ja que, com s'explica en la introducció, les dades de qualitat son un requeriment clau per generar sistemes i decisions de qualitat.

En projectes de mineria de dades i intel·ligència artificial, independentment de l'àmbit de treball, l'esforç i temps que es dedica a la preparació, processat i manipulació de les dades previ al desenvolupament de models està, en general, al voltant del 80% de l'esforç global del projecte, sent aquests els processos que marcaran la qualitat final dels models intel·ligents. Per tant, si les dades d'entrada son bones, aquests temps i esforços es redueixen de forma substancial, amb la corresponent reducció dels costos associats i increment en la qualitat dels resultats.

És important disposar d'un bon equip d'adquisició i gestió que faciliti la captura i emmagatzematge de la informació, amb l'objectiu de definir una base de dades estructurada i eficaç que permeti i simplifiqui els posteriors processos d'anàlítica. Les dades s'han de capturar i emmagatzemar lligades a un instant de temps i a una entitat, ja sigui animal, zona, material, etc., que habitualment s'identifica mitjançant un identificador únic. També és una bona pràctica l'execució de tasques de validació i neteja de dades abans que aquestes siguin emmagatzemades en la base de dades per, de nou, simplificar les posteriors tasques d'anàlítica.

A més, la tecnologia actual permet la integració de sistemes externs com serveis de meteorologia o altres sistemes digitals de proveïdors o clients, lligant d'extrem a extrem tot el flux d'informació relacionat amb la nostra cadena de valor (en el cas que ens ocupa, traçabilitat, condicions i vida de les truges).

La base de dades ha d'assegurar el correcte emmagatzematge de les dades, aquesta ha de ser consistent i robusta, igual que la informació que conté, i incorporar sistemes de seguretat per no perdre dades en cas de fallida. Les dades s'han d'emmagatzemar amb el format adequat, dades numèriques com a números, i no com a text, variables qualitatives correctament discretitzades i, en general, evitar la possibilitat d'emmagatzemar informació com a text lliure, ja que cada usuari s'expressa de forma diferent i l'anàlisi d'aquesta informació és complicada de forma desmesurada. Si no es pot evitar introduir dades de forma manual, es recomanable contractar personal tècnic

qualificat i format en aquesta disciplina que entengui i compregui quins son els requeriments i necessitats que hi ha en passos posteriors respecte la informació capturada.

La captura de la informació s'ha de realitzar de forma correcta i tenint en compte què és el que s'està mesurant, ja que si aquest procés es realitza de forma inadequada, les dades capturades poden resultar inservibles i la inversió realitzada no es podrà justificar. En aquest aspecte, hi trobem dos elements, la precisió i la freqüència. Els sensors han de ser capaços de capturar la informació de l'entorn amb el detall correcte. Per exemple, en els sensors de temperatura aplicats al cas que ens ocupa, una precisió de mig grau és suficient per permetre als sistemes intel·ligents identificar variacions representatives en aquesta mesura. En la mateixa línia, la freqüència en la que és captura la informació també ha de correspondre al context de la variable mesurada. En el cas del pes d'un animal, és considera suficient realitzar una mesura diària per representar la seva evolució al llarg del cicle de procés. Per tant, la selecció del hardware de monitorització a utilitzar es troba directament lligada amb l'element a observar, i és molt recomanable realitzar un estudi o consultar un expert per validar aquesta selecció.

Com a resum, existeixen 5 passos importants per a una correcta recollida de dades i que s'utilitzen en qualsevol àmbit:

1. **Pensar i determinar quina informació es vol recollir:** El primer que es necessita és identificar quines condicions del procés productiu a estudiar tenen afectació en aquest i cal conèixer. Tenint en compte quins objectius es volen aconseguir, s'ha de decidir quines variables s'han de recollir, per proporcionar les dades que permetin assolir-los. El coneixement previ dels experts en el domini ajudarà a aquesta tasca.
2. **Determinar els intervals de temps en la recol·lecció de dades:** En els primers passos de planificació, també s'ha de determinar en quins intervals es recolliran dades. Per exemple, si es vol mesurar el consum animal, la recol·lecció no cal que sigui horària, sinó que amb una periodicitat diària, la informació ja es prou precisa per a la majoria de objectius d'analítica en el sector.
3. **Determinar com es recollirà cada variable:** En aquest pas s'elegeix el factor clau de la recollida de dades, decidir la tecnologia a usar. Per seleccionar el format de recollida de variables, s'ha d'avaluar quin tipus d'informació es recollirà i quina quantitat. És important prioritzar l'automatització en la captura de dades, ja que la recollida manual no sempre és exacta i pot derivar ens problemes d'interpretació. En canvi si la recollida es fa mitjançant sensors, el procés es pot automatitzar i només es requereix un cert manteniment temporalment.
4. **Començar a recollir dades:** Un cop completats els punts anteriors, ja es pot començar a emmagatzemar i utilitzar les dades.
5. **Millora contínua de la recollida de dades, avaluar les dades recollides i la idoneïtat dels anteriors passos:** Quan ja s'està recollint informació, s'ha de seguir avaluant que les dades recollides tenen el format i qualitat esperats i el procés no té fallades, sinó, és necessari revisar els passos anteriors.

Finalment, el correcte manteniment dels sensors i sistemes digitals és un altre element crític, ja que sense aquest, les lectures dels sensors es poden veure afectades per desviacions o averies (per



exemple derives en les senyals) que alterin la realitat i afectin les dades i en conseqüència els resultats de posteriors estudis.

## 4. Perquè cal fer-ho

L'anàlítica de dades habilita la contínua millora productiva en qualsevol procés mitjançant l'estudi dels esdeveniments passats amb l'objectiu de descobrir punts de millora que permetin escalar en la reducció de costos de producció i ampliar la capacitat de generar valor afegit.

En el context del projecte d'anàlítica centrat en el desenvolupament d'un sistema de millora de productivitat en granges reproductives, s'han desenvolupat models de predicció de productivitat de truges d'acord a l'històric de cicles productius. En aquest sentit, en una granja de cria, periòdicament es reemplacen algunes de les truges mares menys productives. Per tal de dirigir aquesta selecció de truges, s'executa un model intel·ligent que prediu la productivitat de les truges presents a la granja en el seu següent cicle reproductiu. Els resultats obtinguts permeten predir amb un fiabilitat del 80% la productivitat d'una truja (que estigui en el seu quart cicle reproductiu) en el següent cicle. Aquests resultats varien en funció del cicle reproductiu de la truja però mitjançant aquestes prediccions i el càlcul d'un índex de productivitat per truja, s'aconsegueix generar una classificació segons la futura productivitat de les truges de la granja. La selecció de les truges més mal puntuades per l'algorisme permet tenir una recomanació objectiva de les truges que cal reemplaçar, amb l'objectiu d'assegurar una qualitat mínima en les truges presents en el cicle de producció.

Com a exemple, i paral·lelament a l'execució del projecte d'anàlítica, al llarg dels processos d'exploració i processament de dades s'han extret conclusions que han permès detectar la necessitat de millorar els processos de recollida de dades del procés productiu. A continuació s'expliquen aquestes conclusions i quines millores hi estan relacionades.

Primer de tot, s'ha comprovat que no totes les causes de baixa poden ser predites, ja que causes com una patologia, la baixa condició física o fins i tot la mort no poden ser predites amb les dades (en aquell moment) disponibles. La possible millora proposada consistiria en afegir variables que identifiquessin l'estat físic actual de la truja, com podrien ser comportaments anòmals, el consum d'aigua i menjar de la truja de manera precisa, anteriors registres de patologies o el pes de l'animal.

En l'estudi realitzat, els registres de truges contenien la data d'entrada de l'animal a la granja, però es desconeixia la seva trajectòria anterior a aquest esdeveniment, com per exemple la data de naixement de l'animal. Aquesta fet fa que el càlcul de l'edat de l'animal, es realitzi de forma aproximada a partir de la data d'entrada, sabent que els porcs podien entrar a una granja amb edats diverses. Aquest càlcul és un exemple de manca d'informació, ja que no es comptava amb dades de l'animal en la seva etapa anterior a l'entrada a la granja, i no es podia discernir entre diferents condicions inicials. Durant l'exploració del procés de gestació i deslletament, s'ha pogut comprovar la importància d'obtenir informació relacionada amb els garrins. Tot i que en general es prefereixen les truges amb parts nombrosos, hi ha la probabilitat que l'estat físic dels garrins sigui pitjor i per tant el risc de mort o el temps que tarda a deslletar sigui més alt. En les dades estudiades, el pes de cada garrinada acostuma a correspondre a una mitjana dels garrins nascuts o una aproximació feta pel

cuidador. Tot i que el pes és una bona indicació, l'equip creu que també és important millorar la notificació de la baixa condició física dels garrins.

Per últim, la importància de variables d'entorn (com poden ser les condicions ambientals o la qualitat dels aliments) és elevada per als problemes de predicció, ja que un algoritme ha de poder tenir en compte les condicions d'entorn on es troba un animal. Si volem aconseguir un algoritme capaç de fer prediccions acurades independentment de la granja estudiada, s'han d'utilitzar variables no només individuals, sinó també de l'entorn de l'animal.